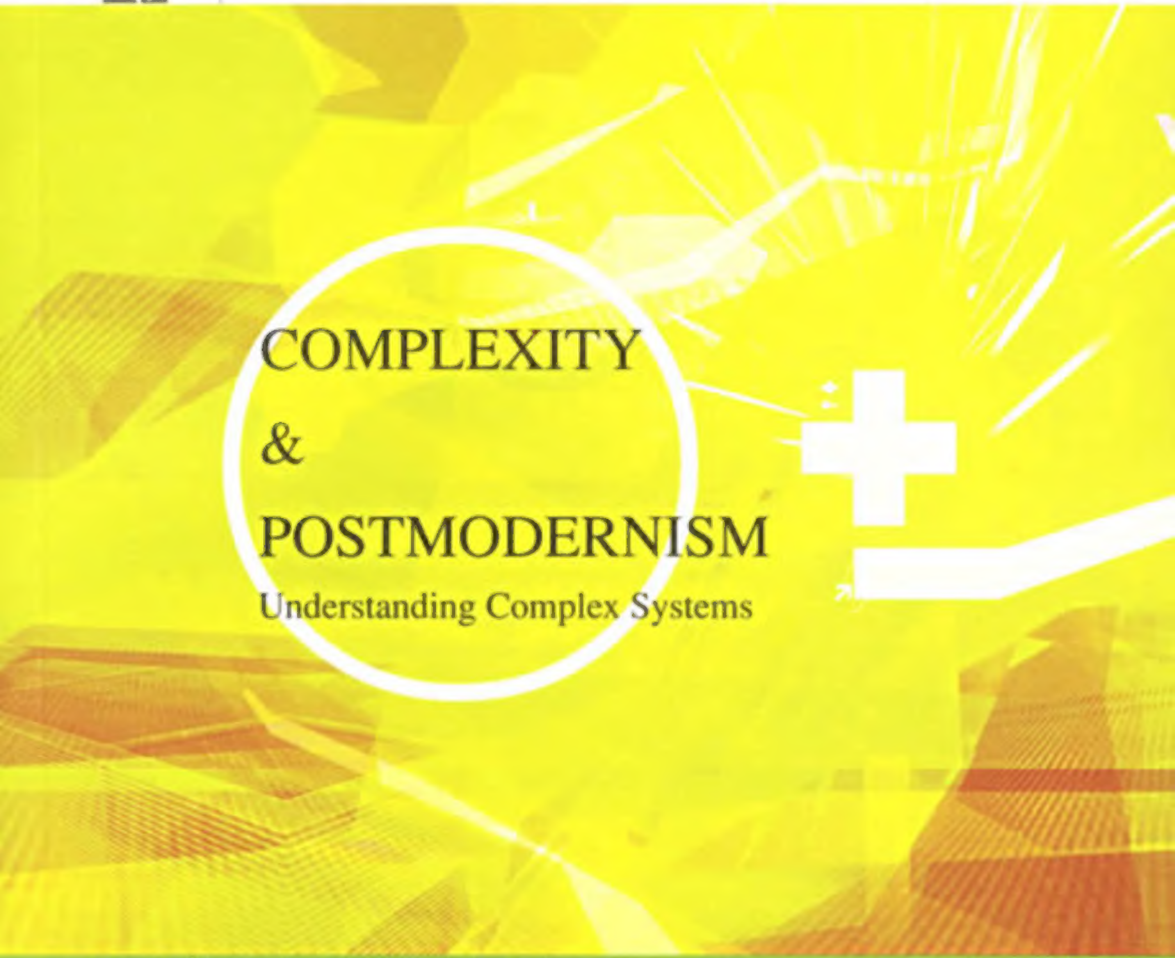
开放人文



[南非] 保罗·西利亚斯 著　　曾国屏 译　　Paul Cilliers

复杂性与后现代主义

理解复杂系统

上海世纪出版集团

世纪出版

ISBN 7-5428-3916-0

9 787542 839169 >

定价：21.00元

易文网：www.ewen.cc

复杂性与后现代主义

理解复杂系统

[南非] 保罗·西利亚斯 著　　曾国屏 译

世纪出版集团　上海科技教育出版社

图书在版编目(CIP)数据

复杂性与后现代主义：理解复杂系统/(南非)西利亚斯(Cilliers, P.)著；曾国屏译. —上海：上海科技教育出版社，2006.4
(世纪人文系列丛书)
ISBN 7-5428-3916-0

Ⅰ.复… Ⅱ.①西…②曾… Ⅲ.科学哲学—研究
Ⅳ.N02

中国版本图书馆 CIP 数据核字(2005)第 077310 号

责任编辑 章 静 姚 宁 潘 涛
装帧设计 陆智昌

复杂性与后现代主义——理解复杂系统
[南非]保罗·西利亚斯 著
曾国屏 译

出　版 世纪出版集团 上海科技教育出版社
(200235 上海冠生园路 393 号 www.ewen.cc)
发　行 上海世纪出版集团发行中心
印　刷 上海宝山江杨印刷厂
开　本 635×965 mm 1/16
印　张 15.25
插　页 4
字　数 183 000
版　次 2006 年 4 月第 1 版
印　次 2006 年 4 月第 1 次印刷
ISBN 7-5428-3916-0/N·677
图　字 09-2004-030 号
定　价 21.00 元

出版说明

自中西文明发生碰撞以来，百余年的中国现代文化建设即无可避免地担负起双重使命。梳理和探究西方文明的根源及脉络，已成为我们理解并提升自身要义的借镜，整理和传承中国文明的传统，更是我们实现并弘扬自身价值的根本。此二者的交汇，乃是塑造现代中国之精神品格的必由进路。世纪出版集团倾力编辑世纪人文系列丛书之宗旨亦在于此。

世纪人文系列丛书包涵“世纪文库”、“世纪前沿”、“袖珍经典”、“大学经典”及“开放人文”五个界面，各成系列，相得益彰。

“厘清西方思想脉络，更新中国学术传统”，为“世纪文库”之编辑指针。文库分为中西两大书系。中学书系由清末民初开始，全面整理中国近现代以来的学术著作，以期为今人反思现代中国的社会和精神处境铺建思考的进阶；西学书系旨在从西方文明的整体进程出发，系统译介自古希腊罗马以降的经典文献，借此展现西方思想传统的生发流变过程，从而为我们返回现代中国之核心问题奠定坚实的文本基础。与之呼应，“世纪前沿”着重关注二战以来全球范围内学术思想的重要论题与最新进展，展示各学科领域的新近成果和当代文化思潮演化的各种向度。“袖珍经典”则以相对简约的形式，收录名家大师们在体裁和风格上独具特色的经典作品，阐幽发微，意趣兼得。

遵循现代人文教育和公民教育的理念，秉承“通达民情，化育人心”的中国传统教育精神，“大学经典”依据中西文明传统的知识谱系及其价值内涵，将人类历史上具有人文内涵的经典作品编辑成为大学教育的基础读本，应时代所需，顺时势所趋，为塑造现代中国人的人文素养、公民意识和国家精神倾力尽心。“开放人文”旨在提供全景式的人文阅读平台，从文学、历史、艺术、科学等多个面向调动读者的阅读愉悦，寓学于乐，寓乐于心，为广大读者陶冶心性，培植情操。

“大学之道，在明明德，在新民，在止于至善”（《大学》）。温古知今，止于至善，是人类得以理解生命价值的人文情怀，亦是文明得以传承和发展的精神契机。欲实现中华民族的伟大复兴，必先培育中华民族的文化精神；由此，我们深知现代中国出版人的职责所在，以我之不懈努力，做一代又一代中国人的文化脊梁。

上海世纪出版集团
世纪人文系列丛书编辑委员会
2005年1月

对本书的评价

这是一本令人信服的原创性重要著作。

——玛丽·赫西(Mary Hesse),
剑桥大学科学哲学荣誉退休教授,
《科学哲学的革命和重建》的作者

这是一本关于复杂性理论的优秀导论，为复杂性理论在心智哲学、科学哲学，以及社会理论方面的意义提供了一个机智的述评。

——基思·安塞尔·皮尔逊(Keith Ansell Pearson),
沃里克大学教授

《复杂性与后现代主义》产生出一种即席而稳健并能够“归纳出答案”的思维意象，这也正是西利亚斯在这本令人着迷、恰逢其时的著作中所展示的调研模式。

——保罗·哈里斯(Paul Harris),
洛约拉·玛丽蒙特大学教授

我已经将此书读了三遍。每次阅读都感受愉悦，收获颇丰。它清晰简明，写作的确是优美的。……这是一本充实了丰富论据的重要著作。书中洋溢着美好的事物。

——戴维·伯恩(David Byrne)，

《人工社会和社会模拟杂志》

作 者 简 介

保罗·西利亚斯(Paul Cilliers)，1956年出生于南非弗里尼欣，先后获电气工程学位和哲学博士学位。在作为研发工程师的10余年中，他专长于利用神经网络进行计算机建模和模式识别。他在斯泰伦博斯大学约翰·德根纳(Johan Degenaar)和剑桥大学玛丽·赫西(Mary Hesse)的指导下，完成博士学位，后被分派到斯泰伦博斯大学哲学系，主要讲授文化哲学与解构(cultural philosophy and deconstruction)，兼授科学哲学和科学伦理学。目前的研究集中于复杂性理论对理解伦理学、法和正义的意义。2000年，西利亚斯因其出色的研究而获得雷克托奖(Rector's Award)。

前　言

“复杂性”（complexity）与“后现代主义”（postmodernism）都是充满争议的概念。当代社会容易被描述成后现代的，但是对此种描述的反响却大相径庭。在一些人［如鲍曼（Zygmunt Bauman）］看来，后现代主义使得我们有可能逃离现代主义的束缚，使世界返魅（re-enchant）。在另一些人［如盖尔纳（Ernest Gellner）］看来，它体现着相对主义，即一种在其中怎样都行（anything goes）的理论框架。这弄得人们晕头转向。后现代主义也可被视作现代主义的寄生物，或是现代主义的软肋。在这种情况下，人们便会主张，如果我们希望超越这种过分简化的启蒙理想，那就应该彻底放弃这个概念。

对于后现代主义有种种不同反响，是以对于这个词的意义有种种不同理解为基础的。即使有可能澄清这种争论，这却不是本书的意图；本书也不试图为后现代主义提供某种辩解。这里主要关心的是复杂性和复杂系统（complex systems）的概念。本书的讨论中，所谓

的后现代主义，只不过是将一些理论趋法*［如德里达（Derrida）和利奥塔尔（Lyotard）的观点］松散地（甚至是不恰当地）归拢在“后现代”的麾下，它们对于所处理现象的复杂性是内含着敏感性的。这些趋法认为，不应以单个或者根本性原则来分析复杂现象，而应承认不可能对真正复杂的事物给出单一的、排他的叙述。不过，承认复杂性，肯定不意味着“怎样都行”的结论。

“复杂性”概念也并非是单义的。首先，有必要将“复杂”（complex）与“复合”（complicated）两个概念加以区分。一个系统，尽管它可以是由极其大量组分构成，但倘若可从其个体组成而获得关于系统的某种完整描述，这样的系统仅仅就是**复合的**。喷气式飞机和计算机就是复合的。另一方面，系统组成之间、系统与环境之间具有相互作用的**复杂**系统，则具有这样的属性：作为整体的系统不可能只通过分析其组分而得到完全理解。而且，这样的一些关系并非固定不变的，而是流动着、变化着，常常是作为自组织（self-organisation）的结果。这会产生出新的特征，通常称作**涌现性质**（emergent properties）。大脑、自然语言和社会系统是复杂的。理解这种复杂性，是贯穿本书的中心问题。

其次，有必要对复杂性和混沌理论（chaos theory）之间的关系作一交代。由混沌理论制造的轰动已经部分地消退了，但是认为它对于复杂系统研究仍具有重要作用的观念则依旧流传甚广。尽管我不

---

* approach（approaches），是一个在科学研究中使用的越来越频繁的词汇。而相应的译法颇多，有译作“探究方法”、“探索方式”，也有译作“进路”、“方式”乃至“方法”等等。按照《新英汉词典》，其基本意思包括：1.靠近，临近；接近，近似；2.进路，入门；3.探讨；处理；态度；方法。这里结合其“接近”、“进路”和“方法”等基本意思，考虑到“方法”的本义是“沿着道路前进”，而采取了“向某某大致目标趋近之法”的理解，从而尝试在此将之译成“趋法”。相应地，对于动名词的 approaching，往往译成了“趋进”。——译者

否认混沌理论对于复杂性研究能够有所贡献，但我确实也感觉到，这种贡献会是极其有限的。例如，在对复杂系统进行分析时，对初始条件的敏感性并不是一个如此重要的问题。事实上，正是复杂系统的**鲁棒**(robust)本性，即其在不同条件下以同样的方式发生作用的能力，保证了系统的生存。蝴蝶翅膀的扇动在地球的另一方引发风暴的比喻，是描述对于初始条件的敏感性的一个很妙的比喻，但是我感觉到它也引起了太多的混乱，应该完全弃而不用。混沌行为(chaotic behaviour)——在“决定论混沌”(deterministic chaos)专门术语的意义上——源于相对少量方程的非线性相互作用。然而，在复杂系统中却总是存在着相互作用着的极其大量的组分。尽管人们获得了关于嗅觉系统或心室纤维颤动的功能性方面的描述，但我却难以相信在自然界中发现的任何行为都能够被描述成该专门术语意义上的真正混沌。我发现，一旦需要在系统的不同状态之间的明显转化，自组织临界性(self-organised criticality)概念(参见第 6 章)较之混沌的比喻要更为恰当。这也许听起来有些过分轻率了，我也的确不想声称，混沌理论(或分形数学)的若干方面不可能有效地运用于关于自然的建模过程。我倒是认为，混沌理论特别是决定论混沌和普适性(universality)概念，并没有真正帮助我们理解复杂系统的动力学。作为分形数学样板的芒德布罗集(Mandelbrot set)，常常被当作我们已知的最复杂的数学对象，但在最终分析的意义上，它是复合的而不是复杂的。在眼下的研究框架中，混沌理论仍作为现代范式(modern paradigm)的一部分，但不会对其细节作深入探讨。

本书的目标，是要从后现代，或者也许更精确地说是从**后结构**(post-structural)的视角来阐明复杂性概念。从这种视角获得的最明显的结论是，不存在总揽一切的、可忽略掉复杂系统中偶发性方面的

复杂性理论。如果某事物的确是复杂的，就不可能从简单理论得到恰当的描述。关注复杂性，必须关注特定的复杂系统。尽管如此，我们还是可以在非常基础的层面上，对复杂行为和复杂系统动力学的条件进行一般性说明。而且我还认为，可以对复杂系统进行建模。模型可以在计算上执行，并可能导致可以执行更复杂任务的机器。不过，模型本身将至少与被建模的系统同样复杂，因此不可能导致我们对于系统本身理解的任何简化。作为这种模型的一个例子，我广泛地使用了神经网络，即一种称作联结论的趋法。事实上本书通过指出神经网络的运行与德里达的语言工作的描述之间具有结构相似性，来理解后现代理论对于复杂性研究的意义。

除了关于联结论(第2章)和后结构主义(第3章)的导论式章节，以及不考虑塞尔(Searle)对于该种争论的贡献(第4章)之外，本书讨论的中心问题是表征(representation)(第5章)和自组织(第6章)。关于表征概念的讨论，或说是关于表征理论的解构，以例子说明了，基本的哲学分析可以对建模技术作出贡献。反过来，对于自组织——一个在科学文本中经常(但肯定不是唯一)遇到的概念——有助于我们获得这样的哲学观点，即一个并非先定或固定结构的系统，其行为不一定是随机的或混沌的，换言之，并非怎样都行(anything does not go)。

本书并不打算系统地探讨道德理论，当然也不可能是在一个价值无涉的空间运行。因此，伦理问题总是会不时浮现，尤其是在第7章。本书中所揭示的复杂性和复杂系统的特点，对于有待发展的社会的、道德的理论，肯定是有启发意义的。我希望，这在进一步的研究工作中成为更为中心的方面。

我乐意借此向促进了本书中思想的形成的以下人士表示感谢：

约翰·德根纳(Johan Degenaar)、玛丽·赫西(Mary Hesse)、詹尼·霍夫迈尔(Jannie Hofmeyr)，以及斯泰伦博斯大学的两个学科际研讨小组，其中一个以文科学院为基础，另一个以理科学院为基础。埃斯玛丽·斯密特(Esmarié Smit)对于手稿的完成提供了难以估量的帮助。

在第2章、第3章和第7章中被使用的一些材料的前一版，已经发表于《南非哲学杂志》(South African Journal of Philosophy)。对于获准在这里再次使用，也要表示感谢。

目录

1 对本书的评价
3 作者简介
1 前言

1 第1章 趋进复杂性
3 1.1 复杂系统的概貌
10 1.2 复杂性的量化
14 1.3 复杂系统不可或缺的两种能力
17 1.4 复杂系统的建模：两种趋法
25 1.5 基于规则的模型与联结论模型的比较
29 1.6 联结论与后现代主义
32 1.7 为复杂性建模

35 第2章 联结论概述
36 2.1 理论背景
43 2.2 语言规则
49 2.3 联结论与后现代科学

51 **第 3 章 后结构主义、联结论和复杂性**
52 3.1 索绪尔的语言模型
57 3.2 仅作为能指系统的语言
63 3.3 网络和系统

66 **第 4 章 塞尔迷雾**
68 4.1 中文房间
72 4.2 中文房间的框架
76 4.3 塞尔对后结构主义的拒斥

81 **第 5 章 表征的一些问题**
83 5.1 真表征
93 5.2 联结论
103 5.3 实际建模中分布式表征的重要性
110 5.4 表征的解构

123 **第 6 章 复杂系统中的自组织**
124 6.1 若干关键方面
126 6.2 自组织系统的属性
132 6.3 自组织临界性
136 6.4 通过选择而组织
145 6.5 自组织的哲学含义

155 **第 7 章 复杂性与后现代主义**
155 7.1 承认复杂性
156 7.2 后现代主义

165 7.3 复杂系统与后现代社会
171 7.4 作为复杂系统的语言
175 7.5 从后现代视角理解科学知识
189 7.6 后现代伦理学的复杂性

196 **第8章 后记 理解复杂性**

199 **参考文献**
214 **译后记**

第 1 章　趋进复杂性

科学的世界和哲学的世界从来就不是彼此孤立存在的，但是人们也许会认为，这两个世界的关系正在进入一个新阶段。普遍存在的进行应用研究的压力，肯定与此有关，但是技术重要性(importance of technology)的巨大提升，却是另一个明显没有那么多政治性的原因。初看起来，人们也许认为这降低了哲学视角的重要性，以为哲学的重要性总是只与理论的重要性相联系，但是我认为，真实的情况恰恰与此相反。这并非是理论不重要，或者科学的理论方面不是哲学性的。如同当今理论物理学那样“哲学”的科学努力已经不多了。科学(以及我们生活的世界的其余部分)的技术化(technologisation)正在根本性地改变着科学与哲学之间的关系。

我们处在这种变化进程之中，要清晰地描述出正在发生着什么并非易事，但是问题的关键在于，我们的技术已经变得比我们的理论更加强大。我们能够去做我们尚未理解的事。我们可以进行基因剪接(gene-splicing)，却尚未完全了解基因是如何相互作用的。我们可以

研制药物，却还无法解释其效用和预见其副作用。我们可以制备出亚原子粒子，却并不精确地知道它们在实验室之外是否实际存在。我们可以贮存、提取无穷无尽比特的信息，却并不知道其含义。所有这些发展，其核心之处是由于电子计算机的表现能力。它构成了我们绝大部分工具（如洗衣机和汽车）中的一部分；它渗透到我们的社会各界（想一想金融业和娱乐业吧）；它迅速地成为了最重要的传播媒介。尽管我们知道在一台计算机中并没有出现什么“陌生”的东西，但是当一台计算机在执行复杂的任务时，却没有人能够完全地把握其中发生的所有方面，至少是无法追溯到在0和1之间的开关转换层。这只是因为太复杂了。

技术的力量（power of technology）已经为科学打开了一种新的可能性。一种最重要的科学工具永远都是分析方法。如果某种事物过分复杂，难以作为一个整体来把握，就把它分为若干份可以分别分析的可操控单元，然后再将它们合在一起。不过，对于复杂动力学系统的研究，已经揭示了此种分析方法的一个根本缺陷。复杂系统并非仅仅是由其组分之和构成，而且也包括了这些组分之间的内在的错综复杂的关系。在“割裂”（cutting up）一个系统时，分析的方法破坏了所要寻求理解的东西。幸运的是，这并不意味着对于复杂性的探索是没有希望的。以强大的计算机为支持的建模技术，允许我们对复杂系统的行为进行模拟，而不必非得理解它们。我们利用技术可以做到我们利用科学无法做到的事情。在过去十年中，对于复杂性理论兴趣的增长，因此也就不奇怪了。

强大技术（powerful technology）的兴起，并非是一种无条件的幸运。我们不得不处理那些我们尚未理解的东西，而这需要有新的思维方式。正是在这种意义上，我认为哲学将发挥重要的作用，这不

是对科学和技术中所发生的事物提供某种元描述(meta-description)，而是作为科学和技术实践中的一个整体的组成部分。一些特定的科学哲学的视角会对我们趋进复杂系统的方式产生影响，我希望指出，其中的一些视角——常常被标记为后现代——对于研究复杂性具有特殊价值。为了将后现代理论的某些方面运用于复杂系统的研究，首先要建立起何为复杂系统的一般性理解。

1.1 复杂系统的概貌

在此，期待可以对“复杂性”意味着什么至少能提出某种工作性定义。不幸的是，无论是定性上还是定量上，该概念都还是难以把握的。卢曼(Luhmann 1985：25)提出了一个有用的描述，他认为，一个系统中，复杂性意味着可能性要多于可实现性。这很难被当作定义，但如果复杂性难以被赋予某个简单的定义，人们也就不要对此感到惊讶。相反，我们可以尝试对复杂系统的特征进行分析，以建立某种一般性描述，而不试图限制于某个特定的先验(a priori)定义。这也就是本节中所要进行的尝试。在下一节我将转向复杂性的量化问题。

在转向讨论复杂系统的某些特征之前，我们必须要注意到两个重要的区别。“简单”与“复杂”之间的区别，并不像我们直觉认为的那样鲜明(Nicolis and Prigogine 1989：5)。许多系统貌似简单，但仔细考察时却显示出显著的复杂性(如树叶)。另外一些貌似复杂，却可以简单地描述，例如某些机器(如内燃机)。对于复合性事物，复杂性并非是定域于系统的某个特定的可辨识的部位。因为复杂性是系统中组分之间的相互作用引起的，所以复杂性展现在系统自身的层面上。无论是在较低的层次上(某种根源)，还是在更高的层次上

(某种元描述)，都不可能捕获复杂性的本质。在复杂和简单之间的这种区别，常常成为我们与系统之间的“距离”的函数(Serra and Zanarini 1990：4，5)，即我们用来描述系统的函数。(从远处看)养鱼缸作为一件装饰品是相当简单的，但是(从近处看)作为一个系统，它却是相当复杂的。这并不意味着复杂性只是一种语言现象，或者只不过是我们描述系统的函数。复杂系统的确具有并不仅仅是由观察者的观点所决定的特征。不过，它的确意味着，谈论复杂性时必须加以小心。简单和复杂往往是互相掩饰的。

第二个重要但同样难以一以贯之的区别，是复杂和复合之间的区别。一些系统，有非常大量的组分，承担着精致复杂的任务，但却能够被精确地加以分析(在该词的完整意义上)。这种系统，是复合的。另一些系统，由错综复杂的非线性关系和反馈回路所构成，每次只可能对某些方面加以分析。而且，这些分析总是会引出种种曲解。这样的系统，是复杂的。曾听人说过(当然，是由某个来自法国的人说的)，喷气式飞机是复合的，而蛋黄酱却是复杂的。另一些在原则上可加以精确描述的系统，如CD播放器、雪花、芒德布罗集，也是复合系统的例子。复杂系统通常是与活事物联系在一起：细菌、大脑、社会系统、语言。这种区别，尤其是被强大的新技术(例如，有非常大的存储器的快速计算机究竟是复杂的还是复合的)所削弱了，但是将它用于建立一种对于复杂系统特征的描述则很有用。我在此提供一张清单[1]：

(i) 复杂系统由大量要素构成。当要素数目相对较小时，要素的行为往往能够以常规的术语赋予正式描述。不过，当要素数目变得充分大时，常规的手段(例如某个微分方程组)不仅变得

不现实，而且也无助于对系统的任何理解。

（ii）大量要素是必要条件，但非充分条件。我们并没有兴趣将海滩上的沙粒当作复杂系统（来研究）。要构成一个复杂系统，要素之间必须有相互作用，而且这种相互作用必定是动力学的。一个复杂系统，会随着时间而变化。这种相互作用，不一定必须是物理的，也可以设想成信息的转移。

（iii）相互作用是相当丰富的，即系统中的任何要素都在影响若干其他要素，并受到其他要素的影响。不过，系统的行为，并不是由与特定要素相联系的相互作用的精确数量所决定。如果系统中有足够的要素（其中有一些冗余），若干稀疏关联的要素也能够发挥与丰富关联的要素相同的功能。

（iv）相互作用自身具有若干重要的特征。首先，相互作用是非线性的。线性要素的大系统通常会崩溃成为小许多的与之相当的系统。非线性也保证了小原因可能导致大结果，反之亦然。这是复杂性的一个先决条件。

（v）相互作用常常是作用于某个相对小的短程范围，即主要是从直接相邻接受信息。长程相互作用并非不可能，而是实践上的制约迫使我们只能作这种考虑。这并不预先排除大范围的影响——因为相互作用是丰富的，从一个要素到任何另一个要素的途径通常包含着若干步骤。结果是，相应的影响也按此方式进行了调整。这可以通过若干方式得以增强、抑制或转换。

（vi）相互作用之间形成了回路。任何活动的效应都可以反馈到其自身，有时是直接的，有时要经过一些干预阶段（intervening stages）。这样的反馈可以是正反馈（加强，激发），也可以是

负反馈(减低,抑制)。两种反馈都是必要的。在复杂系统中相应的术语叫做归复(recurrency)。

(vii) 复杂系统通常是开放系统,即它们与环境发生相互作用。事实上,要界定复杂系统的边界往往是困难的。系统的范围并非系统自身的特征,而常常由对系统的描述目标所决定,因而往往受到观察者位置的影响。这个过程被称作构架(framing)。封闭系统通常都只是复合的。

(viii) 复杂系统在远离平衡态的条件下运行。因此必须有连续不断的能量流保持系统的组织,并保证其存活。平衡不过是死亡的另一种说法。

(ix) 复杂系统具有历史。它们不仅随着时间而演化,而且过去的行为会对现在产生影响。任何对于复杂系统的分析,如果忽视了时间维度就是不完整的,或者至多是对历时过程的共时快照。

(x) 系统中的每一要素对于作为整体系统的行为是无知的,它仅仅对于其可以获得的局域信息作出响应。这一点极其重要。如果每一要素对于作为整体的系统将要发生什么都“知道”,那么所有的复杂性都必定出现在那一要素中。这会导致,在单个要素并不具有必要能力的意义上的物理上的不可能性;或者在某一特定单元中整体的“意识”的意义上,构成了一种形而上学的冲动。复杂性是简单要素的丰富相互作用的结果,这种简单要素仅仅对呈现给它的有限的信息作出响应。当我们观察作为整体的复杂系统的行为时,我们的注意力就从系统的个别要素转移到了系统的复杂结构。复杂性是作为要素之间的相互作用模式的结果而涌现出来[2]。

让我们通过一些例子来说明这些特征吧。以雪花为例。从远处看，雪花是显得非常简单的对象，但是当我们靠近观察，它展示出惊人的细节。雪花是按六角形排列的，六个角的“每枝”都显示出精细而美丽的模式结构。尽管所有的雪花都具有这种形式，但每一片特定的雪花又是有差异的。一片雪花由大量的要素（水分子）构成，要素之间通过其结晶结构而相互作用。每一个分子都仅仅受到局域信息的影响（雪花中分子并非由于外部决定而必须采取此种位置），但分子之间的关系是相当固定的。在此没有真正的反馈回路，也没有进化（也许除了融化之外）。就其结构而言，它不是真正的开放系统。它是一种暂时的平衡，无法去适应环境，因此会迅速地丧失其结构。一片雪花，尽管外观上奇妙复杂，但仅仅是复合的。

接着，让我们考察一些真正的复杂系统。人的大脑被看作是已知的最复杂的对象。类似地，我们每天用来交流的语言，也不能完全有分析地描述。这两种复杂系统——大脑和自然语言——在本书的研讨中将受到深入的关注。因此，我将利用其他的例子——经济系统——来阐明复杂系统的10个特征[3]。

为了架构出我们的描述，我们必须确定与此系统的“距离”，换言之，我们将考虑到何种的细节程度？如果站在很远的地方，也许只能考虑大的金融体系的活动——银行、大型公司，甚至是国家。显然，在此过程许多较小的细节都会丢失。如果准备对系统从微观细节上进行考察，也许必须保持着对每一便士的状态进行监控。在这种情况下，我们便面临着由于较低层次的乱哄哄的活动将所有有意义的模式都模糊掉的危险[4]。为了论证这一点，让我们以如下方式架构出系统：此种方式允许我们考虑个体的人——作为经济行为者(economic agents)的能力——作为我们的复杂系统的要素，并将此系

统的边界划定在一个国家之中。此复杂系统的10个特征将以如下的方式展现出来：

（i）在一个国家中，经济上活动的人肯定包括了大量要素，通常有数百万人。

（ii）形形色色的个体通过贷出、借进、投资和购物而交换货币。这些关系处在不断地变化之中。

（iii）一个经济行为者与大量的其他要素发生作用：商店、银行和其他行为者。一些行为者比另一些行为者更为活跃，但这并非是他们转移的金钱数量的函数，也不标志他们对于系统的影响。

（iv）相互作用是非线性的：金钱可以获得复利，小额投资可能产生大额回报（例如适时购买了恰好的股份，反之亦然）。

（v）经济行为者主要是与其他近邻（不仅指空间意义上）发生作用：地方的商店或服务提供者，以及其同事或合作者。不过，他们能够方便地通过诸如银行或经纪人这样的中介与更远的部分发生作用。

（vi）一个行为者的活动可以最终反射回自身。好的投资可能会有好的回报（正反馈），过度的花费可能导致货币供给的短缺（负反馈）。没有反馈，就没有经济系统——如果没有回报，那么谁会进行投资呢？活动也可能经过了大量的中介步骤之后而反射回来。通货膨胀复杂性即是一个好的例子。

（vii）经济系统必定是开放的。要划出其边界实际上是不可能的。它不断地受到政治系统、农业（以及因此气候条件）、科学和技术、国际关系、社会的稳定性等等的影响。日用品、产品、

货币和信息持续地在系统中流动。

(viii) 因为经济系统是受到供需动力学驱动的，所以永远不可能处于平衡态。它可以增长或收缩，上扬或下降；但它绝不会静止不动，即便是不景气时也不会。甚至当我们提到“稳定的”经济时，“稳定性”也必须以动力学术语来理解。

(ix) 经济系统受其历史的影响极大。今天的价格很大程度上取决于昨天的价格。许多重要的经济趋势的长期变化相当缓慢，但是特定的影响则可能引起急剧变化。

(x) 一个经济行为者只是在存在可利用信息时才能发挥作用。他并不知道所有其他行为者在干什么。例如，当一个行为者希望采购日用品，他的决策是依赖于若干的“局部”因素而作出：我真正想要多少？我买得起吗？可以到别的地方去购买吗？等等。这种对通货膨胀率的效果、支付的平衡、投资者的信心、利率等等的作用，通常都不在考虑之列，即使它的确会影响(微小地，但绝不比其他类似的作用要小)所有这些因素。

我们对经济系统的描述也许看起来微不足道(在许多方面的确如此)，但是对此却有一些好的理由。我们已经描述了系统中的要素及其在运行层面上的相互作用。如果我们想要将此讨论转向更为复杂的经济现象(国民生产总值、股票市场指数、黄金价格等等)，也不需要增加任何其他的东西；这里讲到的经济现象，也正是由于系统中种种要素之间的相互作用而涌现出来的。这些相互作用通常采取了相互合作中的要素成团集簇的形式，并与其他团簇相互合作，也与其他团簇相互竞争。例如，一家银行，不过就是为了实现特定功能而由

一些个别群体集合而成。复杂经济系统的组分并非由不同类型的事物(银行、国家、公司和个人)组成;它们是由个体行为者聚集在一起形成的大尺度现象。我们希望获得理解的其中的高阶复杂性(higher-order complexities),并不存在于任何个别行为者之中,而是存在于它们之间丰富的相互作用的模式之中。

经济系统的例子使得我们最后可以指出非常有意义的一点。系统中的一个要素,可以属于一个以上的团簇。一个人与一个以上的银行打交道,可以在一家大公司工作并自在地游荡于股票市场。团簇不应该只从空间意义上来解释,或者被看作固定的、与外界隔绝的实体。它们可以生长或收缩,可以被分解或吸收,可以繁荣或衰败。团簇是动力学的,并与其他团簇发生着相互作用,这种相互作用既可以是直接的,也可以是通过个体成员相互共享的。

这里提出的10个特征,有助于我们以定性的方式谈论复杂性,但是并没有给我们提供某种度量复杂性的方法。我们希望能够以更定量的方式来处理复杂性,特别是当为复杂系统建模时。这在何种程度上是可能的?

1.2 复杂性的量化

19世纪下半叶热力学的发展,特别是玻尔兹曼(Ludwig Boltzmann)的工作,是最初成功地处理了复杂现象的尝试[5]。通过表述出三个"定律",科学家能够以精确的方式处理能量的利用和转移,而不必卷入到低层次的复杂性中。

> ……平衡态热力学是物理学对自然界的复杂性问题作出的第一个响应。这个响应是用能量的耗散、初始条件的忘却、趋向无序

的演化这样一些术语来表达的。经典力学，即关于永恒和可逆的轨道的科学，不同于被进化概念统治的十九世纪所面临的那些问题。平衡态热力学能够把它自己关于时间的观点(即认为时间隐含着衰退和死亡的观点)与其他科学的观点对立起来。(Prigogine and Stengers 1984：129) *

经典力学中，时间是可逆的，因此并非是方程的一部分。在热力学中，时间扮演着一个关键性角色。这也许在热力学第二定律中有最好的表述，即一个系统的熵只可能增加。熵可以被视作系统"无序"的度量。随着系统转化能量，可利用的能量越来越少，系统中"无序"在增加。

熵是一个复杂的概念，香农(Claude Shannon)以其天才的创意，用它来度量消息的信息量。在两篇开创性的论文中(Shannon 1948, 1949)，他建立了一种关于通信的数学理论，形成了现代信息理论的基础。通过将热力学方程中的"能量"代之以"信息"，他表明，消息中的信息量等于其"熵"。一条消息愈是无序，其信息量便愈大。可以用一个例子来说明这种论证。考虑在一次传给你一条由数字串组成的消息，例如，如果数字串仅仅由一系列3组成，你将会在接收端迅速地注意到这一点。接下去的数字是可预测的，它将没有任何新的信息。尽管消息是高度有结构的，但其信息量却是非常低。序列1489 1489 1489的信息量要稍高一点，但是也会很快变得可预测。下一序列中的数字对于接收者愈是难以预测，也就是消息

* 译文参见：伊·普里戈金，伊·斯唐热.从混沌到有序——人与自然的新对话.曾庆宏，沈小峰译.上海：上海译文出版社.1987.173——译者

中的信息量愈大。一条大信息量的消息，便是一条对其结构可预测性低的消息，因此“熵”也大。

因为香农的理论漂亮地形式化了，它就可能方便地运用于许多工程问题。尤其是通讯技术，得益极大，他的工作的重要性在此是没有非议的。不过，信息的熵理论，其中有一个成问题的含义：如果信息等于熵，那么具有高信息量的消息便是完全随机的消息。显然，在“信息”与“随机”两个概念之间有某种平衡。尽管熵理论(entropy theory)是优美的，但仍然有理由认为，它对于理解诸如人的认知这样的复杂系统(参见 Katz and Dorfman 1992：167)并非是一个合适的模型，在此，错综复杂的结构肯定不能等同于“随机”。

由于蔡廷(Gregory Chaitin 1975，1987)的工作，一些有趣的新视角被提了出来。在他命名为“算法信息理论”的关于信息理论的再诠释中，随机性不是根据不可预测性，而是按照“不可压缩性”来定义。他关于随机性的定义如下：“一系列的数字是随机的，如果最小的算法能够以与系列相等的信息比特数将其向计算机进行说明。”(Chaitin 1975：48)[6]考虑一显著长(比如说长度为1000)的数字序列。如果这个系列仅仅由3组成，显然就可以写出非常简单的计算机程序将其生成。比如：

第一步：打印“3”。

第二步：重复第一步 1000 次。

此程序明显要比最初的序列短得多，因此具有非常低水平的随机性。随着数字的序列变得更加复杂(!)，要生成出它的程序的长度必然变

得更长。当程序变得与序列一样长时，那么就说序列变得随机了。

对于一定的序列，显然可以有许许多多种程序，其中的一些可以是相当精致和冗长的。我们仅仅关心能够执行此任务的最短或最小程序。这样的程序被称为最小程序(Chaitin 1975：49)。注意，最小程序按照定义自身是随机的，而无论它生成的此序列是随机的或不是随机的，因为它不可能被进一步压缩。因此，有可能将任何非随机序列归约成随机序列(记住，程序——在理论的数字上——只不过是数字的序列)，而此随机序列是一个较短的序列。

通过这种过程，随机性成为序列中信息量的度量，但是，极其重要之处在于，随机性不再从不可预测性进行理解，而是从信息密集的密度来进行理解。这还向我们提供了一种有趣的关于复杂性的定义：一条序列的复杂性，等于需要产生出该序列的最小程序的大小(Chaitin 1975：49)[7]。

在一定的意义上，我们经过一番努力终于获得一个自明之理：复杂性就是复杂(complexity is complex)。如果一个复杂系统不能简化到从一个简单系统(或者也许并非只是复合的)为起点，那么这个复杂系统就不能被归约成一个简单系统。这种见解，对于许多科学家所持有的一种理想是有启示的：去发现那支配着自然中一切事物的基本原理。分析方法的成功，已经创造了一种幻觉，即认为所有的现象都受到一组能够明晰化的定律或规律的支配。奥卡姆剃刀(Occam's razor)，常常是在调查研究的初期就启用，就体现着这种信念。蔡廷的分析有助于使我们认识到，真正的复杂性问题只可能以复杂的资源进行趋进。这种认识，也是一种对于反还原主义见解的重新诠释。它并不否认，复杂系统是以通常的材质组建起来的。不过，它的确否认，对这些组分及其相互作用的描述，实际上就可以解释系统作为

整体的行为。一个复杂的系统不可能还原成其基本构成组分的集合，并不是因为此种系统不是由它们组成的，而是因为在此过程中会丢失过多的关系信息。

最后的一些评注，涉及蔡廷关于复杂性的定义：蔡廷的工作，尽管处在非常低的层次上，但却是在数论形式系统的语境中进行的。我认为，这样的形式系统，对于给复杂性进行一般性建模，提供的是一个不适当的起点。严格说来，他的分析至多可以适用于复合的现象。然而，在直觉的水平上，“不可压缩性”概念仍然是非常富有成效的。它提示我们，当处理复杂性时，不存在没有风险的捷径。然而，此概念不应该绝对地加以运用。我们感兴趣的复杂系统，绝非是彻底“最小的”；它们包含了大量的备用能力或冗余。对此种必要性，至少有一种以上的理由：它提供着鲁棒性、发展的空间，以及弹性的手段。至于为何某种“自由空间”是必要的，在下一节中考察我们的模型必须要抓住的复杂系统的两个重要方面时，更清晰的图像将产生出来。眼下我们可以得出结论，对于复杂性的严格度量，看来是不可能的。为了描述复杂系统，在一定的意义上，必须重复此系统。

1.3 复杂系统不可或缺的两种能力

复杂系统必须应付变化着的环境。相应于这些变化的激烈程度，其对系统的资源将有大的需求。为了应付这些需求，系统必须具有两种能力：必须能够贮存关涉环境的信息以备未来之运用，以及必须在必要时能够适应性地改变其结构。其中第一种能力将作为表征过程加以讨论；第二种能力涉及在没有外部设计者的先验必然情况下内部结构的发展和变化，将作为自组织过程加以讨论。

这些能力至关重要，将独立成章详细论述。在本节中，它们仅仅是作为宽泛的术语而被引入。不过，重要的是要认识到这些能力对于我们复杂系统的模型的意义。真正复杂系统的任何模型，都将必须具备这样的能力。换言之，表征和自组织过程必须在模型中就可加以模拟。这意味着，必须给予这两种能力以某种描述形式。我将指出，这是可能的。

1.3.1 表征信息

为了对环境作出适当的反应，一个复杂系统必须能够收集起关于环境的信息，并贮备起来以供未来运用。换一种表述法，系统的结构不可能是由要素的随机集合组成；它们必定具有某种意义。在传统的哲学术语中，这意味着，系统必须要“表征”出对于其存在具有重要性的信息。

对这种表征过程的机制的追寻，造成了一组长久的哲学问题。大脑是如何表征世界的？语言组分与其描述的对象之间的关系是什么？理论何时应对其试图解释的现象进行适当描述？对于这些问题已有一些解答，但是通常都只假定了系统要素与特定外部原因之间的一一关联。这种原子论趋法(atomistic approach)是分析方法的遗产，通常采取将系统结构拆分，将结构的意义置于独立的层次上。例如，在语言中，将语言的结构(语法)与语法单元的意义(语义)作了划分。通常将这两个层次当作相互独立的。语法层次被看作是对于语义层次的特定补充。一个概念，可以用西班牙语，也可以用日语来“执行”，但仍然保持着其本质意义。类似地，世界上的一个对象，可以在大脑中，也可以在计算机中加以表征——在此，执行不同，但表征相同。

我认为，对于描述复杂系统，关于表征的这种理解是不适当的。

意义不是由符号与某个外部概念或对象的一一关联所赋予的，而是由系统自身的结构组分之间的关系所赋予的。这并不否认在系统外部和内部之间的因果关系。然而，它的确否认系统的结构是由外部所决定的。意义是过程的结果，而这种过程是辩证的——涉及内部和外部的要素。这种过程也是历史的，因为系统先前的状态是至关重要的。这种过程，发生在活跃的、开放的复杂系统中。

在专门讨论表征问题的一章（第 5 章）中，我将提出“分布式表征”概念。在此种框架中，系统的要素自身并不具有表征的意义，而仅仅是与许多其他要素的关系的模式。而且，意义的抽象层次（“语义”层次）成为多余的，或不如说，对于系统的建模过程是不必要的。分布式表征，在联结论的（或神经）网络中可得到最好的执行，而且我将指出，这些网络为复杂系统提供了合适的模型。在更仔细地描述这些模型之前，与传统的趋法相比较，我们必须提到复杂系统的另一种能力，即发展起有组织的结构的能力。

1.3.2　自组织

诸如一个活的有机体或一个发展中的经济这样的复杂系统必定要发展其自身的结构，并能够适应其结构以应对环境的变化。声称某种外部设计者（或上帝）不过是在回避复杂性起源的问题，而我们必须找到系统能够获得并适应其在进化基础上的内部结构的机制。

在此关键性的概念是自组织。这并不意味着，复杂系统包含着某种形式的内部“主体”，由此控制着系统的行为；事实上，整个的中心控制的概念都变得可疑了。本书中将要讨论的自组织（第 6 章）是一个过程，系统藉此可以从颇为无结构的起点发展起某种复杂结构。这个过程，改变了系统在外部环境和系统历史两者的影响下，其分布式要素之间的关系。由于系统必须应对环境中的不可预测的

变化，系统结构的发展也就不可能包含在某种控制着系统行为的刚性程序中。系统必须是“可塑的”。它还表明，自组织过程可以从数学上建模。“自生”(autopoiesis)概念，如同“涌现”(emergence)概念一样，并不涉及任何神秘的东西。

对复杂系统的一般性描述，将受到复杂性模型的影响。这些模型必须共享它们所建模型的系统的特征。接下去，我们要讨论两种可能的模型。

1.4 复杂系统的建模：两种趋法

1.4.1 关于模型的两种视角

我们为什么要为复杂性建模呢？可以从两种视角来回答。从传统的科学观点看，为了预测和控制复杂系统的行为，就需要模型。其优点是明显的。较好的模型可以使科学家更牢固地理解诸如在经济学、生物学、医学、心理学、社会学、法律和政治中遇到的复杂性。只有对于疾病、精神疾病、犯罪和变化无常的经济系统有了更好的理论解释，我们才能从中受益。不过，为了有效果，这些模型必须起作用，必须产生出结果。然而，这里还有一些问题。如何对这些模型进行检测和评价？它们的精确性应该有多大？细节上应该考虑到何种程度？大体说来，关于复杂性的特定科学模型都有一些严重的局限。最成功的模型，是其局限性清晰可见的模型；保持在约束之中的模型，是由特定的建模过程所决定的。这里列举的模型将包括流体动力学和湍流、统计量子力学，在计算机科学中设计良好的专家系统以及信号处理系统。

从更一般的哲学观点来看，可以说，我们希望对复杂系统建模，因为我们希望更好地理解它们。对于我们的模型相应最主要的要求

是，从必须是正确的转移到必须是信息丰富的。这并不意味着模型与系统自身之间的关系变得不那么重要了，而是意味着，从控制和预测转移到理解，的确对于我们趋进复杂性是有效的：对模型的评价是看它能否顺从地起作用。一旦我们对于复杂性动力学有了更好的理解，我们就能够寻求在不同复杂系统之间的相似和差异，并因此建立起对于不同模型的优势和局限的更清晰理解。这种观点还允许我们进行更多的推测，并将那些从严格的科学观点看来是不允许的思想结合进来。

这种研究的目的之一在于表明，为了给复杂性建模，我们既需要科学的视角，也需要哲学的视角。概而言之，人们可以说，科学没有哲学是盲目的，哲学没有科学是无效的。它们之间的合作将使双方获益。一方面，一旦我们开始对复杂性的本质有了更多的理解，那么关于复杂性的模型才在科学实践中变得成功。另一方面，只有将我们的模型进行实际的科学测试，才有可能改进我们的理解。因此，学科际的趋法能够为研究开辟新的途径。例如，它可以表明，(来自传统的哲学领域中)一定的语言模型是等价于一定的大脑模型的，而后者一般被看作是属于自然科学的领域(参见 Cilliers 1990)。随着我们对于复杂性的探索的继续，从一个领域的发现推论出另一领域的知识将得到强调。

在本节的最后，我将讨论关于"我们如何对复杂性建模"问题的两种可能答案。两种趋法都可以按照科学术语来建模，都具有其自身的哲学传承。

1.4.2 基于规则的符号系统

在过去的 40 年中，对于人工智能(简称 AI)的探索是复杂性模型研究的聚焦点。数字计算机能力的迅速增长，使得人们期望的计算

机具有智能行为将成为可能。这样的智能可以达到何种程度（次于或是超过人），取决于你对于此种 AI 的方法和工具的力量的信念。一般说来，这种研究的早期阶段（直到 20 世纪 60 年代初）是以大量的乐观主义为标志的。很快，计算机解决了数学问题，还可以下棋，但是智能的重要标志——感知、运动和语言运用——的复杂性超过了所有的估计。在这些方面，没有任何计算机装置的能力可以哪怕只是遥远地接近人的能力。

尽管对于 AI 能够办到什么的期望已经有了一些降低（而所需要的时间尺度则增加了），但如果宣称此纲领失败了却是不正确的。沿着此方向，产生出许多相当有用的副产品。专家系统就是一个例子。这些系统将特定的知识领域（如地质勘测、宫颈癌诊断）作为一组计算机的规则，将某个“专家”的知识进行了建模。非专家随后得以在特定的领域中利用这种系统来执行某些任务。尽管专家系统处理了复杂性和智能的某些方面，但是它们是以僵硬不变而著称的。本研究将讨论 AI 范式的缺点问题；不过，首先我们需要对形式符号系统的基本特征加以描述。

一个形式系统，是由一些记号或符号组成的，如同游戏中的棋子。这些符号可通过一组定义了允许什么和不允许什么的规则（例如象棋规则）结合成模式。这些规则是严格的形式的，即它们遵从严格的逻辑。在任一特定时刻，符号的构型组成了系统的一种“状态”。一种特定的状态将激活一组应用规则，从而将系统从一种状态转移成另一种状态。如果支配着该系统行为的一组规则是精确的和完整的，人们就可以检验系统的各种各样的可能状态是否是允许的[8]。

我们现在有了一组具有一定构型的符号。下一步是使得这些符

号“表征”某种事物。例如，如果其中的每一符号代表语言中的一个语词，那么系统的规则(语法)可以决定(形式)语言中使用语词的种种组合。于是，系统允许的状态翻译成了该语言的有效语句[9]。对符号的解释，也被称作系统的“语义学”，是独立于支配此系统的规则的。因此，假如这些系统能够表明在逻辑层面上相互等价的话，同一个语义层次就能够在不同的实际形式系统中执行。

形式系统可以是非常简单的，例如，井字棋(tic-tac-toe)中所需要的棋子和规则；也可以是极其错综复杂的，例如，一台现代的数字计算机。计算机能够胜任的事的范围，向我们显示了形式系统的威力。这个范围是如此地令人眼花缭乱，使得我们往往忘记了它是一种形式系统，仅仅由一组按规则操作的记号组成(在此称作程序)。计算装置的形式方面，在被称作图灵机的自动形式系统的抽象模型中得到了最好的描述。从这些数学模型，计算机的一般性和威力变得清楚了。人们还可以表明，在不同类型的图灵机和不同的形式语言如乔姆斯基(Chomsky 1957)所运用的形式语言之间的等价性。

基于符号规则的系统，构成了关于复杂性建模的经典趋法。复杂系统的行为必须归约成一组适当描述该系统的规则。问题就取决于找到那些规则，这里假定了它们是存在的。一个如同自然语言或人脑的系统能够真正归约成一组规则吗？经典的 AI 主张是能够办到的。在转向另一种趋法之前，让我们总结一下基于规则的系统的主要特征(依照 Serra and Zanarini 1990：26)：

- 基于规则的符号系统是在抽象的(语义的)水平上对复杂系统进行建模。符号被用来直接地表征重要概念。以此方式，复杂系统的许多偶发方面，即运行的不必要的细节，都可以忽略掉。

模型由该符号之间的一组逻辑关系(生成规则)构成。

- 这组规则受到中心控制系统的支配,被称为系统的元规则。这种控制系统决定着在计算的每一阶段何种生成规则应该被激活。如果中心控制失效,那么整个系统也就失效了。
- 每一概念都有一个专示符号,或者相反地,每一符号都表征了一个特定概念。这被称作局域表征。表征理论,对于精神和语言的形式解释处于核心地位,如同可以在乔姆斯基/福多尔(Fodor)模型中所见。

在这里已经小结的,是常常运用于AI研究的一般方法。这种方法显然不是唯一的方法。研究人员在许多细节上不尽相同,并非每人都主张基于规则的系统具有同样的范围和威力。一般认为,在强AI支持者与弱AI支持者之间有分歧,强AI认为形式系统为人工智能的所有方面提供了适当的模型,而弱AI则仅仅把形式系统看作一种强有力的工具。在第4章中,当我们分析塞尔(John Searle)反对强人工智能的论据时,其中的一些张力将得到阐明。这种案例研究,还对将形式系统用作大脑和语言的模型作进一步阐述。我认为基于规则的系统不是复杂系统的合适模型的观点,主要依赖于我关于局域表征的批判(第5章)。

1.4.3 联结论模型

在讨论复杂性时,大脑总是占有一个特殊位置,这不仅仅在于其自身的结构复杂性,而且还因为它处理复杂性的能力。例如,大脑是如何执行如运用语言、演奏提琴这样的复杂任务的?人们期待着,像这种问题可以促进对于大脑自身进行建模的努力。

像这样的问题,肯定激励了年轻的弗洛伊德(Sigmund Freud)在

1895 年构造出一种相当吸引人的大脑模型(Freud 1950)。该模型不仅奠定了他后来的大部分命题的基础，而且还已经主张了与现代神经学大部分相通的论据(Pribram and Gill 1976)。当麦卡洛克和皮茨(McCulloch and Pitts 1943)建立了一种神经元模型，并表明神经元如何组合起来执行计算时，他们就把关于大脑的建模又推进了一步。这些模型激发了强烈的研究欲望，在 50 年代末罗森布拉特(Frank Rosenblatt 1958)因提出感知机模型获得了可观的成功。这些模型仍然有一些局限性，正如明斯基(Minsky)和帕佩特(Papert)(1969)毫不客气地指出的那样。他们的著作导致了从神经激发的模型转向了形式符号系统。只是在 80 年代初期找到了克服感知机局限的数学方法之后，神经模型才重新获得了它们失去了的基础。从那时以后，这些模型在广阔的应用领域获得了前所未有的关注(Rumelhart and McClelland 1986)。

在本节中，我将引入神经激发的“计算”模型，它们有不同的名称如神经网络、分布式处理器或联结论模型。对其特征、运用和哲学意蕴的详细讨论，将在随后的研究中进行。第 2 章将介绍它们是如何工作的。

从严格的功能观点看，大脑不过就是由丰富的相互关联的神经元所构成的大网络。每一个神经元都可以被看作一个简单的处理器，计算着其所有输入之和，以及当输入之和超过一定阈值，便产生出一个输出。而这反过来成为向所有与此神经元相联结的神经元的输入。每一联结都以突触为中介。突触可以引起进入的信号或兴奋或抑制目标神经元，它也决定着影响的强度。例如，从感官进入的信息以这种方式被处理，并分配到大脑的其他部位，在此可以有特定的效应，例如引起肌肉的运动[10]。

大脑运作的这一层次，可以用相互联结的节点网络来建立模型。

每一节点接受其输入之和，并产生输出。输出由节点的转移函数所决定，必定是非线性的。两个节点之间的联结（“突触”）具有某种“权重”，可以是正的或负的，由此决定着节点A对于节点B的影响的强度。在任何特定的联结中，信息流只有一个方向，但是对于A和B之间的两种联结——一个从A到B而另一个从B到A——是不会被制止的。任何节点还都可以与自身联结，这种联结可以是直接的，也可以是通过其他节点的。

这种网络是如何处理信息的？如果某些节点成为了输入节点，即从网络的外部接受信息，而另一些节点成为输出节点，那么此节点可以“处理”输入，产生输出。输出值由两方面来决定：输入值和网络中已有的权重值。例如，输入可以是来自感光器“视网膜”，而网络的输出可以是联结到灯泡。借助网络中正确的权重值，系统可以感觉到是黑暗还是光亮，并相应地使灯泡亮一些或暗一些。具有不同权重（以及感受器）的此网络，能够实施各种其他的任务，包括识别出变化和趋势。

于是现在清楚了，网络的特征由权重所决定。关键问题如下：不同的权重值从何而来？当然它们可以是由外部行为者设定的，如形式系统的编程员，但是我们希望找到一种系统模型，在此不需要外部设计者，即是可以自组织的系统模型。弗洛伊德在他的大脑模型中（参见 Cilliers 1990）已经详细地建议了此种机制，但是只要涉及我们今天知道的联结论模型，由赫布（Donald Hebb 1949）阐述的使用原理（use-principle）就提供了线索。

赫布提出，在两个神经元之间的联结强度随着其被使用的频度而增加。考虑3个神经元，A、B和C。每次A和B被同时激活，它们的联结强度（W_{ab}）应该略为增加，但是当它们未被激活，W_{ab}便

会慢慢降低。以这种方式，如果 A 和 B 经常被同时激活，W_{ab}将增加，但是如果 A 和 B 仅仅是表面的联合，而 A 和 C 是更常规地联合，那么 W_{ab}将降低，而 W_{ac}将增加。以这种方式，网络中将发展起来内部结构，而仅仅依赖于每一神经元可利用的局域信息。这种结构的发展，也可以叫做“学习”。

为了澄清赫布规则——它已经变得有名了——是如何工作的，让我们返回到上述关于网络的例子，使之必须在天黑时接通灯泡。此网络最初未受到训练，即网络中的权重值是随机的，是不可能执行此种功能的。为了训练此网络，必须每次天黑时将灯泡接通(借助外部行为者)。当灯泡接通，网络中的输出神经元将被迫激活，这将自动地联系着相应的输入，或当没有输入时联系着相应的网络。一旦天亮，就将灯泡关掉，网络现在必须将一个不同的输出条件与不同组的输入值相联系。如果此循环反复多次，网络将会调节其内部权重，而无需外部的干预。通过把赫布规则运用于每一神经元，代表黑暗的输入条件联系着一个激活的输出，反之亦然。网络经过了训练之后，自己便能够执行所需完成的任务。同样地，这一原理可以用来“教会”网络执行完全不同的任务，例如，当土地变得干燥时把水龙头打开，或者识别某人的语音。

以这种方式表述的赫布规则显然只是一种定性的原理，但是已有一些作者赋予它以数学表述(例如 Grossberg 1982，1988)。曾经阻碍神经网络模型发展的主要原因在于，缺乏能够调节处于网络中某处的神经元权重的数学模型，并且没有与输入和输出直接关联。在 20 世纪 70 年代，由于一些独立研究人员的努力使这个问题得到了解决，其中最著名的公式被称作“广义德尔塔规则”(Rumelhart and McClelland 1986，Vol. 1：318～362)。该规则使用“后向传播”方

法来训练前向神经网络(第 2 章中有更详细的讨论)。

我们可以将关于联结论模型的介绍小结如下。相互关联的神经元网络(可以从数学上建模),可以学习执行复杂的任务,无论是通过成功执行了这种任务的实例,还是通过运用内部标准于网络而标志其成功,都表明了这一点。这种任务包括模式识别、发动机控制、信息处理、调节、预测和复制。所需要的只是,在此有某种“感受器”为网络获得信息,以及有某种“发动机”允许输出具有某种外部效应。在网络自身之中,发生的只是神经元以可利用的局域信息而调节其权重。在个体的神经元的水平上没有复杂行为可言,但是神经元系统却能够执行特定的复杂任务。复杂行为从许多简单过程的相互作用中涌现出来,这些简单过程以非线性的方式对局域信息进行响应。

1.5 基于规则的模型与联结论模型的比较

在前一节提到的关于复杂性的两种趋法都有强的支持。采用基于规则的趋法的,有 AI 研究人员、遵循乔姆斯基传统的计算语言学家和认知科学家,特别是那些坚持心智的表征理论的科学家(Fodor 1975, 1981; Sterelny 1990)。支持联结论的,是界定不太明确的、更为学科际的神经科学家、心理学家和工程师群体。有一些工作试图把联结论归结为简单错误(例如 Fodor and Pylyshyn 1988)[11],或是将其同化(Lloyd 1989),或将其部分吸收(Sterelny 1990)成为基于规则的范式。我并不否认,在此有交叠的区域,或者实用的复杂性的模型也许可以将两种范式的若干方面结合起来。不过,为了恰当地理解这两种趋法,我们必须从一开始就聚焦于其差异上。如下的大部分差异,对于反对将联结论同化成为基于规则的范式,是有充分说服力的[12]:

- 不同于形式系统将推理规则应用于逻辑变量，神经网络将进化原理应用于数字变量。不是计算某种解，网络设置成为某种满足了所受约束的条件。
- 神经网中没有经典意义上的中心控制。处理是分布遍及整个网络的，种种组分（或组分组）的作用是动态变化着的。这并不排除网络中的某部分发展起某种调节功能，但是这将由系统的进化需求所决定。
- 基于规则的系统中的每一符号都具有精确的、先定的意义——这构成了局域表征。联结论网络中个体神经元却没有先定的意义。变化着的遍及若干节点的活动模式执行着有意义的功能。这通常被看作是分布式表征。
- 形式系统有界定清楚的终止条件，只有当这些条件得到满足时才产生出结果。联结论系统则趋向于动态地收敛于某个解，通常以渐近方式来实现。这种过程不需要进行终止；事实上，通常它不会达到某种单一的、最终的结论。
- 联结论网络的内部结构通过自组织过程而发展起来，基于规则的系统则必须通过搜索预编程的选择性，而这大致上以一种先验的方式定义了结构。在此意义上，学习是神经网络的一种隐含特征。在基于规则的系统中，学习只有通过明晰表述的程序才可能发生。
- 除了形式的基于规则的模型必须在语义学层次上进行诠释的事实之外，模型本身也分为两个层次：符号的层次和规则的层次；或者是，数据的层次和程序的层次。在联结论模型中只有一个层次，即神经元及其权重的层次。不是“主动的”程序和“被动的”数据，在此是通过相互作用而不断调节的数字权重。

没有任何程序，有的只是记忆。

对于习惯了进行高层抽象并乐于逻辑推理的人们，联结论趋法常常显得含糊而浅薄，过分贫乏。结果是，他们对于此种趋法持严重的怀疑态度。科伦(Kolen)和戈埃尔(Goel)(1991)的工作，就流露出三方面的疑虑：

- 联结论网络不可能表征"……较高阶的关系。这种表征的贫乏导致了缺乏对较高阶进行概括的能力，因为网络只能学习它可以表征的东西"(365页)。
- 联结论学习方法是"弱的"，因为它们不会利用所要进行建模的对象的外部的或先验的知识。领域知识必须由设计者"硬性链接"到网络之中。
- "联结论的学习方法并不反映所论问题的结构"(369页)。同样的学习方法，可用以解决迥然不同的任务。

第一条异议只不过是叙述了对于强表征理论的承诺。网络的确并不"表征较高阶关系"，但是，如果要坚持进行表征，才会是一个问题。钱德拉塞卡朗(Chandrasekaran)等人(1988)对此进行了明确的承诺。在他们看来，抽象的"信息处理"层次高于任何特定的实现，无论是由符号还是由联结论来实现(30页)。正是在此抽象层次上，具有了"解释威力"(33页)。如同福多尔和佩利夏因(1988)以及劳埃德(Lloyd 1989)的主张，联结论保持着对于表征的承诺，纵然这种表征是"分布式"的事实也无关紧要。我将详细地指出(第5章)，分布式表征造成了所有的差异；事实上，它破坏了整个的表征

概念。联结论网络“不能表征”的事实，变成了一个明显的优势。

第二条异议反映了，符号建模者亟需将被建模的范围简约到有限数目的、运用逻辑推理原理明确表述的原理。我们已经指出，当处理真正的复杂性时，这往往是不可能的。联结论模型可以执行复杂性的诸方面，而不需这种简约。那是它们的长处。声称网络中的权重必须由设计者设置在特定值的观点，是不正确的。在大多数情况下，网络中的权重在开始学习之前明显是随机组合的。人们不可能利用先验的范围知识，因为人们通常并不知道这将会涉及范围的某方面。这也大体上回答了第三种异议，即认为联结论模型过于一般化，没有反映出问题的“结构”。结构不可能得到反映，这完全是由于它不可能以符号术语加以明晰。同一个网络可以被教会去执行“迥然不同”的任务这个事实，不是一种弱点，而是此种趋法的威力的一个标志。

塞拉(Serra)和扎纳里尼(Zanarini)(1990：28～29)指出了若干进一步的优点，可以概括为如下两点：

- 网络自然地处理大量的数据。网络的分布式本性，在此方面等价于某种全息图。并不需要有独立的对于每一比特数据的局域表征。输入数据也无须按照符号关系来“诠释”；它们能够以相当“粗糙”的形式送进网络。
- 因为网络以满足约束为基础进行运行，它们可以处理矛盾问题。(基于规则的系统本质上更“脆弱”，遇上矛盾的信息时变得完全不可靠。)这并不意味着，网络能够从“不正确”的数据产生出“正确”的答案，而是说，它不像基于逻辑规则的系统那样会将其放大[13]。一个网络，可以是通过赋予其他约束更多的权重来解决这种矛盾，或是将产生出没有意义的结果。这是联结

论趋法的鲁棒性的一个标志。

作为最后的依据，形式符号系统的支持者可以主张，没有任何事情可以阻挡他/她不断地将特设规则(*ad hoc* rules)加入到系统中，直到它可以执行所需任务。这也许是相当可能的，但是当处理复杂性时，只有在时间和资源可能变成天文数字的意义上，采用这种趋法的主张才是可能的。而且，特设规则越多，模型就变得越是“分布式”。要遵循这种途径来获得一个工作的模型，最终也只可能是停留在极其大量的构建联结模型的方式上。

从稍微更哲学化的观点来看，克拉克(Clark)和卢茨(Lutz)(1992：12)指出了联结论趋法的另外两个优点：

> 首先，我们有一种方法论，外部世界可藉此以更直接的、非特设的方式推动着计算模型。其次，我们有一种系统，其中智能行为不需以表征世界状态的类语句结构的任何推论性推理的准逻辑过程为基础。因此，智能和具有准语言结构的推理得以分离开来。

从这些考虑中，我的观点已经明确：联结论模型对于理解复杂系统以及为其建模，要比基于规则的模型更有用。我希望，这种主张可能在以后的章节中得到具体说明。

1.6 联结论与后现代主义

本研究的目的之一在于表明，划入宽泛的(而且常常是贴错标签的)“后现代主义”范畴中的某些理论的某些方面，对于复杂性的研究具有重要的启发含义。这些含义将在后面的章节中进行更详细的

探讨。这里只是作了一个简介。

尽管利奥塔尔的《后现代状况》(Postmodern Condition)将成为第7章中讨论的基础，但也许更准确地说，焦点将是复杂性理论和后结构理论之间的密切关系。缘由在于，我希望避开那些可能被解释成相对主义的后现代趋法。中心的论据将来自德里达(Jacques Derida 1973，1976，1978)(主要是早期的)工作，无知者往往将这些工作看作是“相对主义的”。

后结构分析(post-structural analyses)主要是运用于语言、文学、文化和艺术，理论的视野却因此更为宽阔。它深深地涉及我们对于整个生活世界的理解和解释，并因此在我们关于科学的理解上具有启发含义。不过，在科学哲学的语境中，它很少被提到，科学共同体对这些发展也就几乎未予关注。如果我们考虑到如下的两个缘由，这种状态也许并不值得惊讶：

- 科学取得了巨大的进展，在20世纪上半叶尤其如此，而其本质上是现代主义的。在理论的层面上，科学与那些结合逻辑的哲学分支如逻辑经验主义及其变迁[通过波普尔(Popper)，拉卡托斯(Lakatos)和库恩(Kuhn)]之间建立了一种强大的联盟。在英国的分析哲学以及在美国的实用主义语境中，科学哲学都是在科学家或多或少颇感舒服之中发展起来的。他们认为这反映了他们的方法和目标。由于这种分析式趋法对于实际的科学实践其实没有什么意义，科学家因此通常是可以忽略哲学的，或者至少是感觉到并不需要必须研究可替代的哲学探索。
- 后结构主义(解构)常常以反科学术语出现，它们强调了意义的多重性，现存等级秩序的打破，逻辑的缺点，分析式趋法的失败。

这种颠覆性的态度，被看作是解构的，是抛弃所有形式的理性，因此就否认了科学所依赖的基础。人们不能由此指责科学家对于以这种方式(在我看来是不正确的)呈现的趋法的质疑。

上面提到的第一个问题，源于分析哲学和大陆哲学之间的一堆差别。这应该在两种传统之间的富有意义的相互作用层面上来加以探讨。第二个问题，由流行于后结构理论中的思维和写作风格所致。为了吸引科学家的更严肃关注，过度热心的后结构主义者(特别是文学理论家)不得不将他们的浪漫色彩转变成某种更冷静、更明晰的东西，某种更易引起争论的东西。一些后结构主义者也许会主张，这种趋法在"真正的"后结构主义中是不可能的；但这不过是一种托辞。我认为，后结构主义并非仅仅是一种话语分析的颠覆形式，它也是一种敏感于所论现象的复杂性的思维风格。由于科学可以从这种趋法中受益，所以相关的思想必须要使得科学家能够接受。受益将包括如下方面：

- 将分析趋法狭义地聚焦在还原主义倾向上是贫乏的。证伪原理仅仅提供了一种摆脱思想的机制。相反，后结构主义是一种包容性趋法，特别是通过跨学科的相互作用，将积极鼓励产生出新颖思想。不同学科中的模型可以转换和结合起来，从而增加在追求特定问题的解答时可利用的选择。
- 后结构主义的趋法将导致一种更多聚焦于实践结果——以及这些结果的含义——的科学哲学，而不那么关注于产生出使得科学知识具有合法性所必备的抽象元叙事(meta-narrative)。
- 一般的科学"方法"，可以用某种对所论问题偶发性

> (contingencies)更为敏感的东西来代替。为了处理常常是极其大量的观察和实验数据，科学传统上极其强调的是遵循正确的方法。实验被设计出来控制大量的变量，并限制结果的可能解释。尽管这种步骤通常是不可避免的，但它意味着一些可能的结果被先验地消除了。选择一种方法就是一种预先腾空，以指向特定的一组解。遵循严格的方法，一定会提供绝妙的结果，但是经常导致没有充分重视对于方法的选取，而且还导致了使用一般术语来解释实验结果，而不是在合适的方法框架之中来解释。后结构主义具有更多的"游戏"的趋法，但是这种态度并非是幼稚或轻浮。处理复杂现象时，绝不是单一的方法将产生出整个的真理。游戏般地趋进复杂系统，允许不同的进路、不同的观点，以及也许能对其特征有更好的理解。

尽管我将强调复杂性趋法的后结构主义和联结论之间有密切关系，但是其中没有任何一个主要的论据预先采取了后结构主义立场。因此，每一章的主题都将在没有明确地参考后结构主义的情况下引入并加以讨论，甚至那些熟悉它们的人也应该承认这种一般策略。对于那些不太熟悉后结构理论的人，第3章中就德里达的思想的相关方面进行了介绍。每一章的结尾会对相应的哲学意蕴作一扼要的讨论。最后一章将对此详细讨论。

1.7 为复杂性建模

为建立起对复杂性本性的敏感性，本章中已经考察了复杂系统的某些特征。为复杂系统建模，即便是建立一个非常有局限的模型，这种敏感性也是必要的。一些要点可以小结如下：

- 对复杂系统的特征更深入的考察清楚地表明，分析方法在处理复杂性时具有局限性。分析式趋法并没有错，但是通过将系统的部分“去掉”来“分析”复杂系统将总是会附加上严重的扭曲。
- 对于复杂性的算法的和信息理论的趋法在其试图揭示复杂性的真正本质时失败了，但这却给我们提供了非常有价值的洞见，即复杂性是“不可压缩的”。一个复杂系统不可能被归约成更简单系统，除非它从一开始就并非是真正复杂的。一个复杂系统模型必将“保留”该系统自身的复杂性。由于复杂系统的模型将如同它所建模的系统般复杂，因此不可能按照一些逻辑原理来揭示出此系统的“真正本性”。
- 复杂系统，其处理信息的方式以及内部结构的发展和变化，都与其涉及的环境具有特殊的关系。
- 计算机技术已经为复杂系统建模开拓了新的可能性。传统的趋法是按照一组从形式符号系统导出的逻辑规则为系统建模。我提呈了一种联结论模型作为替代，并讨论了这两种趋法的差异。
- 一个建议是，在联结论模型与后结构理论之间有着有趣的联系。对这些联系的探讨将贯彻于本研究中。

总之，我将试图再一次推动关于复杂性的联结论模型的选取。复杂系统的一些特征在本章的前面已经指出。复杂系统的模型必须要把握住这些特征。以形式符号系统为基础的模型中，这些特征必须要明确地加以建模。在算法描述中必须要描述高层次上的相互关联性、归复性、分布性等等。在联结论模型中，这些特征都已是模型结构的隐含部分。联结论模型由大量的单元构

成，在这些单元中存在具有反馈回路的丰富的相互关联，但仅对局域信息作出响应。为了具体说明这种观点，下一章中将对神经网络作一稍微专门的介绍。

注 释：

1. 这些特点取自 Nicolis and Prigogine (1989)、Serra and Zanarini (1990) 以及 Jen (1990) 的类似描述。

2. 使用"涌现"概念，通常会让人产生当"事物结合起来"就出现某种神秘事物的印象。在这里的使用意味着没有什么东西是不应该说的。它只是强调了这样的事实，没有什么"外在"，没有外在目的或设计者，它需要"引起"系统的复杂行为。也许使用"关系性质"(relational properties)一词比使用"涌现性质"要更好一些。

3. 选取这个例子是为了说教。我对于经济学没有任何深入的认识。

4. 在一定的意义上，经济系统处在"最低层次"的钱币层次上。没有什么东西更低层，也没有什么东西在高层。经济是由钱币成群结队运动而构成。在德里达的术语中(后面将讨论)，钱币是经济系统的"痕迹"。

5. 对于热力学的发展和意义，及其为研究开拓了新道路的详细讨论，参见 Prigogine and Stengers (1984)。

6. 仅仅是利用一系列数字也许可以看见一点表面，但是要记住，任何形式化都可以通过哥德尔数论系统表达成一系列数字。

7. 要证明一个序列不是随机序列并不难，只要找到一个较短的程序即可。甚至不必是最小程序。然而，要证明一个序列是随机序列，就必须证明不存在更短的程序。蔡廷 (Chaitin) 从哥德尔不完备性定理出发，指出这样的证明是不可能找到的。随机性是可证伪的，但却是不可证实的（这是一个在数学中比经验科学中要强得多的要求）。蔡廷的结论，涉及数形式系统的局限性的论据之含义，这既有趣也重要，但是本研究中不涉及这个问题。进一步的细节可参见 Chaitin (1975：50～53；1987：146～164)。

8. 系统所有状态的状况，不可能由系统内部决定。形式系统的"不完全性"，或换言之"停机问题"(halting problem)，是一个复杂的元数学问题，哥德尔(1962)和图灵(1936)都分别进行过分析，本研究中不涉及这个问题。关于这些问题的通俗性讨论，可参见 Hofstadter (1980) 或 Penrose (1989)。

9. 自然语言是否实际上可以由形式系统来建立模型，是现代语言学中悬而未决的大问题之一。我认为，基于规则的系统对于自然语言的复杂性是不恰当的模型。这里提呈的形式语言的例子，取自 AI 的视角。

10. 这当然是对大脑的实际工作的一种过度简化，但甚至这样的一个简洁的模型，也已经具有了复杂行为的能力。

11. 他们的评论在第 2 章中讨论。

12. 改编自 Serra and Zanarini (1990：26，27) 以及 Chandrasekaran *et al*. (1988)。

13. 在一阶逻辑中，任何东西都可从矛盾的前提出发，而得到证明。一旦出现了矛盾，逻辑推演的结果就变得无意义了。

第 2 章　联结论概述

我认为，复杂系统的基于规则的、分析的传统趋法是有缺陷的，从后现代和后结构理论获得的见解可以帮助我们找到更新颖的方式来看待复杂性。不过，有关的论据并非只处在纯粹理论的层面上。这里还尝试表明，这些见解可以影响我们关于复杂系统的模型。这意味着，建模的“分布式”方法具有复杂系统的某些特征，因此比强表征理论的、基于规则的模型更有前途。

神经网络，或称联结论模型（认知科学家更愿意采用这一名称），并非是唯一的“分布式”建模技术。遗传算法（genetic algorithms）乃至细胞自动机（cellular automata）都具有类似的特征。神经网由于其巨大的灵活性而特别适合，这一点也许是事实，且正是这种考虑影响了将它们作为分布式模型的范例。

神经网络对于许多学科具有重要意义。例如，在认知科学中，联结论引起了一场波澜，有人（Horgan and Tienson 1987：97）将其看作是一场库恩式危机（Kuhnian crisis）的信号。联结论的广泛的哲学

意蕴在这场挑战中得到了强调，它为人工智能研究，更一般地为我们理解大脑、心智和语言，提供了某些基本假设。这种建模技术对于理解复杂系统的重要性，能够以如下方式加以概括：

- 神经网络保持了所要建模的系统的复杂性，因为它们自身具有复杂的结构。
- 神经网络以分布形式对关于其环境的信息进行编码。分布式表征的概念削弱了我们对于传统的表征理论的理解。
- 神经网络具有自组织其内部结构的能力。

其中的后两点将在第 5 章和第 6 章中详加讨论。在本章中，我希望对联结论进行一般性介绍，并将此趋法置于哲学框架之中。这里以联结论的系统对于某些语言能力进行建模为例。本章的中心部分是，考察从分析的视野对于联结论的一个重要批评（Fodor and Pylyshyn 1988）。在对此批评的讨论中，关心的重点是规则在语言活动和心智活动中的作用。从后现代视野来看联结论的可能性，这里只是简单提及，更详细的讨论见介绍后结构主义的下一章。

2.1 理论背景

联结论是由我们对于大脑的理解而提出来的一种信息处理方法。从功能上讲，神经系统仅仅由神经元构成。这些细胞通过突触而大量相互关联。突触传递着在上一个神经元对于排列在下一个神经元的树突所产生的刺激。如果这种刺激超过了一定的阈值，神经元被激发，一个脉冲传导到神经元的轴突。该脉冲反过来向一些神经元提供了突触输入。从一个神经元传导到下一个神经元的信息，被突

触的转移特征以及接受神经元的树突的物理结构所改变。任何单个神经元从许多其他神经元接收输入或向它们提供输出。神经元激发的复杂模式表现为大脑活动的基本特征。

神经元的简单数学模型可以这样来构建。一个神经单元运用其输入之和，决定产生出何种的输出。不过，每一个输入都首先受到一定的“权重”值的增值。这种权重决定着两个特定单元之间的关联强度，也模拟着神经系统中的树突的特征。对任何特定神经元（记作 A）的输出的响应，因此以如下方式计算：所有的关联着 A 的神经元的输出是——在每种情况下都由与特定神经元和 A 之间联结有关的权重所增值——加和在一起的。这个和又被 A 的转移函数对于产生 A 的输出所增值。这种输出变成了网络中的下一个神经元的一个输入，它反过来由途径之中的权重值所调节。此权重值可以为正也可以为负。此转移函数（在所有非平凡的情况下）是一种非线性函数。神经元形成了具有复杂关联模式的大网络的部分，而且因为权重决定着一个神经元对另一个神经元的影响，网络的特征主要由这些权重值所决定[1]。

在网络中，每一个神经元都平行于其他神经元而在连续地计算其输出，由这些权重值所决定的活动模式在此网络中流动。网络的拓扑，即其中神经元关联的方式，也是重要的。一个网络，可以是稀散地关联着、丰富地关联着或完全地关联着。一个完全地关联着的网络，是其中所有的神经元都跟所有其他神经元关联的系统。这种构型被称作归复网络（recurrent net）。归复网络是极其强大的，但是难以加以利用。实践中的神经网络的建构都要简单得多，通常只是“若干层”神经元。图 2.1 中显示了一个简单的神经网络。这是通常称作多层感知机的一个例子。信息提

供给输入层，网络的计算结果可以在输出层获得。在这两层之间，是一层或多层的“隐含”层。它们与外界世界之间没有关联，但是形成了输入和输出之间的联结。

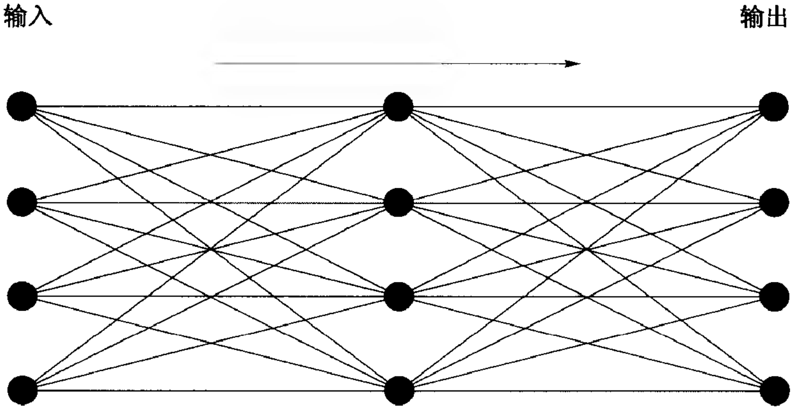


图 2.1　一个简单的前向神经网络

在此网络中，信息仅仅从输入流向输出，神经元按层排列。每一个神经元都与下一层的所有神经元相关联。其中权重没有明晰地表示出来，但是每一线条都具有某种相伴的权重。在这个简单的例子中，同层的神经元之间没有任何关联，其中也没有反馈回路。然而，任何特定神经元的活动都受到其他许多神经元的影响，反之它又对其他许多神经元有作用。信息因此并不定域在网络中的特定部位，而是分布在大量的单元之中。网络的特征由两个权重层决定。

我们现在能够表明，这种网络尽管如此简单，却能够处理信息。如果输入层的神经元以某种方式被激活，随着输入值通过两个权重层的增值，一定的模式将由输出层产生出来。当然，输入和输出意味着某种东西。在此例子中，我们将在稍后见到，输入代表了英语动词的现在时态，输出代表其过去时，是作为网络计算的结果。如果向网络输入动词，权重的模式产生的输出即代表(相应于网络的)其过

去时态。在此关键性的问题是：这些权重，或者至少是这些权重值，是从何而来的？它们可以是由网络的设计者先定的，但是这样的先定，只有对于最简单的问题和最简单的网络才是可能的。然而，对于我们来说，预先知道权重值则是不必要的，因为向网络提供了充分多需要解决的问题的例子，网络将由自身产生出权重值。神经网络是训练出来的。它按照解决问题的方向发生"进化"。

这一点是如何办到的？通过向网络提供输入模式和正确的输出，它能够以使得两个模式更为接近的方式来调节自身的权重。如果在充分长的时间中提供了足够多的例子，而网络又复杂到能够应付这样的问题，一组权重就将会自动地产生出来，使得对于每一输入都产生出适当的输出。在网络经历了训练之后，它不仅能识别那些已经教过的例子，而且对于未知的输入，它还将进行有根据的推测。因为高度的反馈，归复式网络的训练是困难的。对于分层的馈向网络(结构上较为简单，但计算上威力仍然大)，在此有一种以减少每一神经元的局部错误为基础的程序，在假定网络的结构有充分能力去发现解时(即网络特别是其隐含层由充分多的神经元组成)，它保证了向解的收敛(应该有一个解)。这被称作后向传播程序。

让我们指出其中若干至关重要的特征，从而对联结论模型的基本理论进行一个小结。一个由大量简单的神经元组成的神经网络，其神经元有丰富的关联。伴随着这种神经元之间的关联而生的权重，决定了网络的特征。在训练期间，网络会调节相互关联的权重值。任何特定权重的值，都是没有什么意义的；而整个系统中的权重值的模式，则承载着信息。由于这些模式是复杂的，是由网络自身产生出来的(通过可以运用于整个网络的一般学习策略)，因此不存在可用来描述网络解决问题过程的抽象程序。在此存在的，只是复杂的关

系模式。

神经网络是如何执行某些一般但相当复杂的任务的，以下的一些例子将体现出它们各方面的能力。

- 甚至如图2.1所示的简单网络，也能够执行分类任务。考虑若干不同的种类，每一种类都有其自己的成员。网络中的输出神经元的每一个都可以代表一个特定的种类。在训练阶段，每一种类的不同样本作为网络的输入。对于每一特定的输入，与适当种类的输出相联系着的神经元接通。以这种方式，网络学习将一个种类的所有成员都与某一特定输出联系起来。如果训练样本代表着形形色色的种类，那么网络将能够进行可靠的分类，甚至样本并非此训练集中的部分亦如此。

  像这样的分类任务通常称作“模式识别”。一个例子是，网络可以依据其发出的噪声而识别不同种类的运输工具。例如，当使用一组机动自行车、摩托车和卡车产生的噪声谱训练这种网络时，网络将学习如何将它们区别开。如果训练的样本是每一种类的代表，网络随后能够将特定类型的摩托车划分为摩托车，甚至对于以前未曾遇到过的类型也能作出这种划分。网络并没有发展起实施此识别任务的特定算法；它只是通过网络增值了输入的信息。在训练阶段发展起来的正确的权重构型，产生出所需的输出。此网络也可以重新加以训练，以执行截然不同的任务。

- 网络可以实施映射(mapping)。一定的输入构型可以转变成一定的输出构型，这是分类能力的推广。在此情况下，网络的输出并不由一个仅仅代表了某一特定种类的激活神经元组成，而

是遍及整个输出层的激活的模式。映射的本性仅仅受到网络复杂性所制约。如果网络充分地复杂，就可以施行任意的映射(Hornick *et al*. 1989)。下面得到更详尽讨论的将动词的现在时转变成其过去时的网络，就是这种映射的一个例子。

- 一个可以实施曲线拟合(curve-fitting)的网络。考虑如下形式的函数：$y = f(x_1, x_2, x_3 \ldots x_n)$。如果输入神经元取值 x_1 至 x_n，输出神经元的近似值为 y，如果网络充分地复杂到近似于函数 f。网络被教会对于给定的 x 值获得正确的 y 值。如果训练的样本具有代表性，网络将根据一组输入值精确地"拟合"而得到一条近似于 f 的曲线。值得注意的是，这里并不需要知道 f 是什么。对于一组 x 和相应的 y，网络将产生一组权重去自动地拟合曲线。不过，对此不要惊讶，这些网络在内插法方面要好于外插法。这些网络常常用来执行控制功能，以及用作预测(如用于股票市场的)分析。
- 神经网络可以实施自动聚类(clustering)。对于一个两维的神经元阵列，其中每一神经元都接受同样一组输入值，该网络将聚类成输入矢量$(x_1, x_2 \ldots x_n)$，其中通过激活聚集的神经元而聚集起来。这是通过内部的引起网络发生内部结构自组织的竞争过程而实现的。第6章中对于神经网络的这种能力进行了详细的分析。

神经网络解决问题的方式具有某种哲学重要性。为了说清楚这一点，我将把联结论置于更一般的理论框架中。可以鉴别出两种范式，每一种都联系着一位著名的语言学家的名字——一位是乔姆斯基，另一位是索绪尔(Ferdinand de Saussure)。这样的两分法总是有些人为性和简单化，但是对于将信息处理风格与一定的智识传统联系

起来则非常有用。这对于研究联结论与当代大陆理论之间的联系特别富有成果，因为它允许我们从不同的视角来探讨后现代理论。表2.1中总结了这两种范式。

表2.1 两种信息处理范式

乔姆斯基	索绪尔
语言和心智模型，按照生成规则系统工作，并以形式和表征术语进行解释（福多尔）	语言和大脑模型，按照关系系统工作，不能以表征术语进行理解（弗洛伊德、拉康、德里达）
图灵机提供的数学模型	联结论网络提供的数学模型
制作出的机器：数字计算机	制作出的机器：神经网络
“智能”应用：专家系统	“智能”应用：模式识别

乔姆斯基范式所坚持的方式是众所周知的。乔姆斯基运用生成规则的形式语法对语言进行描述，与图灵（Alan Turing）建立的数学模型是一致的。这些模型被称作图灵机[2]，也为数字计算机的描述提供了数学基础。普特南（Hilary Putnam）提出图灵机是恰当的大脑模型［这个理论被称作功能主义，后来普特南（1988）不再支持这一点］，而福多尔（Jerry Fodor 1975）把基于规则的语言系统与心智的功能作用结合起来，不顾执行采用的是何种硬件，从而扩展了乔姆斯基的理性主义程序。这种范式形成了认知科学和人工智能（AI）研究的基础。对于此种范式所能期待的只是智能行为应该被描述为遵循规则的，由此我们有了被称作专家系统的基于规则的AI模拟。

以索绪尔冠名的范式，常常并不与计算理论联系在一起。一个缘由在于这样的事实，一般而言的认知科学和特定的联结论是在那并不熟悉索绪尔原理的语境中得到实际运用的。本书的部分目的也就是要建立起对于此种范式的敏感性。特别是当索绪尔理论的表述方式被诸如德里达（在第3章有更详细的讨论）这样的后结构主义思想家

批评时，对于语言的索绪尔描述与大脑工作的方式之间的关系得以建立（在索绪尔的描述中，语言中每一要素只有在它区别于该语言中所有其他要素的程度上才有意义）。人们有可能利用神经网络理论对索绪尔的语言学提供数学模型，这等价于形式语法提供了对乔姆斯基语言学的数学描述。这里我主要关心的是联结论作为复杂系统的一般模型，以及我还将试图表明，在联结论模型与传统的模型之间有着重要的区别。这些区别在很大程度上取决于规则的作用和状态，特别是涉及语言时尤其如此。

2.2 语言规则

作为其开创性的联结论研究纲领的一部分，鲁姆哈特（David Rumelhart）和麦克莱兰（James McClelland）建立了一种神经网络，可从英语动词的现在时生成过去时（Rumelhart and McClelland 1986, Vol. 2：216～271）。他们没有建立起大范围的语言系统，而是选择了一个受到充分限制的可操控的语言范围，但仍然充分丰富到允许他们论证其主张。生成过去时态的动词，通常是以一定数量的规则来描述的，而且还会存在一定数量的不规则。

他们使用的网络，是一种类似于前面所描述的简单馈向网络。其输入和输出层都由460个神经元组成（239页）。在学习阶段，英语动词现在时的音位表征提呈给网络的输入层，而在输出层得到过去时。在每一种表征中，网络的权重是利用称作“感知机收敛程序”（225页）——也被称作后向传播——的过程来调节的。由网络对于每一输入所生成的输出，在整个的学习阶段不断受到监控。业已发现，网络在此权重采集过程中捕获了包括规则和非规则动词的大部分特征，而且网络可以对在先前未曾遇到过的动词作出适当的反应。

而且，在训练阶段，网络的运行类似于小孩获得过去时态的方式(219页，240页)。一个小孩，最初仅仅知道少量的经常使用的过去时动词。其中大多数是不规则的，但却被正确地使用着。在第二个阶段，网络开始注意到某些模式，出现了过度规则化的过程。使用的动词越来越多，但是以前正确使用的不规则动词现在也开始规则化了。在第三阶段，网络开始注意到了这种差异，规则形式和不规则形式被允许共存。鲁姆哈特和麦克莱兰的网络也模仿了这三个阶段。仅仅利用了联想模式的网络表现出来的这种性能，使他们得出如下的结论：

> 我们相信，我们对于小孩学会在明晰的句子中用英语过去时的表达，已经提供了一种完全不同的见解。我们已经表明，可以不需要借助“规则”作为语言描述的概念，而可以合理地解释过去时态的获得。(267页)

这就是说，尽管规则对于描述语言现象非常有用，但在语言获得或者运用时，并不必须利用明示的规则。这些离经叛道的主张，肯定在乔姆斯基阵营中产生了反响。接下去，我们将考察一下福多尔和佩利夏因是如何驳斥联结论作为一种认知和语言模型的。

在福多尔和佩利夏因(1988)加入讨论的语境中，联结论系统与符号系统形成了鲜明的对照。在符号系统中，信息是由借助于规则而形成结构并进行操作的离散符号来表征。符号系统的本性是由符号所表征的东西所决定的，而不是由系统被操作的方式来决定。如果一个生物系统和一个数字计算机以相同的方式执行着同样的符号，它们在功能上就是等价的。

福多尔和佩利夏因认定了“关于心智的现代理论”的两种传统。他们将一种称作“表征主义的”(Representationalist)传统，而另一种称作“消除主义的”(Eliminativist)传统。

> 表征主义者坚持假定，表征的(或“意向性的”或“语义的”)状态对于认知理论是本质性的；按照表征主义，存在着对于世界状态进行编码的心智状态。与此相反，消除主义者认为，心理学理论可以免除使用这种语义概念作为表征。按照消除主义，对于心理学理论的合适词汇或是神经学的，或是行为学的，或是句法学的；在任何事件中，都不是表征着对象的词汇来作为心智状态的特征。(Fodor and Pylyshyn 1988：7)

十分奇怪的是，他们然后声称，“联结论的建模与表征主义是一致的”(8页)，并由此也站在划分为“经典的”认知理论的方面，即他们所捍卫的立场。在此问题上，它们之间的唯一区别是，经典论者将语义内容规定成“表达”(expressions)，而联结论者则将它规定成网络中的节点(nodes)。他们因此主张，“表征”并非是所争论的问题。那么问题是什么呢？他们是这样总结的：

> 经典理论和联结论理论关于心智表征的本性的见解是不一致的：对于前者而不是后者来说，心智表征展示了一种组合组分的结构和组合的语义学的特征。经典理论和联结论理论还在关于心智过程的本性上见解不一致：对于前者而不是后者来说，心智过程以敏感于它们操作的表征的组合结构为特征。(32页)

这两种差异的结果是什么？一个表征系统以符号来工作。要使符号具有意义，它们首先必须形成结构，其次它必须有可能进行操作。不相关符号的集合是没有意义的。联结论系统是表征的，但是因为它们仅仅是“联想的”表征，也就没有操作它们的规则，不可能用它们来为心智建模。

> 联结论构造的深刻错误在于：因为它承认，心智表征中既无句法结构也无语义结构，所以不能将它们看作是某种生成集，而要看作某种清单。但是清单本身是没有结构的；任何对项目集合的罗列都是一个可能的清单。而且按照联结论原理，任何（偶尔也有关联）表征状态的集合相应地都是一种可能的心智。所以，就联结论构造而言，没有任何东西可以阻止人为的非对称的心智。但是，那种结果是荒谬的。认知能力是结构关联的聚类的结果；它们的系统性（systematicity）是遍及整体的。（49 页）

他们的反驳肯定是成立的，如果人们承认系统是以原子式表征而建立起来的。不过，他们认为联结论系统也适合于那种描述的观点，则是非常错误的。他们在坚持认为信息由网络的节点表征时遗漏了神经网络中分布式表征的真正意义，斯莫伦斯基（Smolensky 1987：137）、德西克（Derthick 1990：255，257）和德什泰尔（Dechtel 1987：22）也指出了这一点。福多尔和佩利夏因选取来体现联结论网络的例子，根本就不是神经网络。它是一个语义网络，一项传统 AI 中众所周知的技术，仅仅描述了一些原子式表征之间的若干关系[3]。他们承认了分布式表征的存在，但是他们仅仅将它当作局域表征的“切片”（sliced thinner）。不过，要点在于：在一个充分发展的

神经网络中，任何节点都没有特定的意义。正如早先已经解释过的，意义取决于权重的值；尤其重要的是，它并不处于任何特定权重甚至是权重群的值之中，而是处在它们每次的关联和激活之中。信息并非贮存于符号之中，也不是由符号所表征，并当必要时（如同在传统的认知模型中）被提取出来，信息是在每次网络的那部分被激活时重新建构起来的（Bechtel 1987：22）。

斯莫伦斯基（1987）也许会论证道，分布式表征是某种形式的表征，但是它与经典的模型是如此不同，以致它们不能置于同一个标题之下。两种系统之间的差异比福多尔和佩利夏因所承认的要根本得多，证据必须作用于某个不同的层次。在这里，人们可以说，在眼下用来展示联结论原理或解决特定问题的有限网络中，斯莫伦斯基是正确的。这些网络针对的是已说清楚了的问题，换言之，是已经勾画出框架的问题。作为一个模型，它们的确代表着问题域，即使是以某种非约定的方式。当我们走进大系统——此时的框架没有那么明显了——时，我确信心智表征的概念此时也不起作用了。在一定的意义上，如果处于“分布式表征”的语境中，则“表征”一词的意义就失效了。世界中的对象，并非是在大脑中“表征的”，正如在自然语言中词并不“表征”某一特定的意义一样。这种主张将在第 5 章中进行讨论。

把此表征的问题搁置一旁，人们当然可以提出这样的问题：是否联结论趋法与基于规则的趋法是互相排斥的；是否其中之一可以归结为另一种。福多尔和佩利夏因（1988：64）承认联结论构造的生物学合理性，以及神经网络可以应用于执行图灵机的事实。他们提出，这是一种可能的选择，然后就敦促联结论者将其研究引导向认真执行“经典”构建[4]。然而，这样的一种策略否认了在局域表征和完全分布式系统之间的重要

差异。斯莫伦斯基(1987：137～143)提出，以“软性”联结论为一方，以“硬性”符号系统为另一方，有5种可能的处理方式：

- 否认其中一方而仅仅继续另一方。对于“软性”的否认也被称作“理性主义”。对“硬性”的否认(按照斯莫伦斯基)导致了例如H·L·德赖弗斯(H. L. Dreyfus)和S·E·德赖弗斯(S. E. Dreyfus)(1986)的“直觉”趋法。
- 允许两种系统作为“处理器”彼此相邻而“共居”。以这种方式，它成为裂脑理论的极端表述。
- 通过模糊逻辑(fuzzy logic)软化“硬性”趋法。这仅仅是模糊了边缘，并使软性成为硬性的量度。
- 使得在本质上“硬性”的系统充分复杂，从而出现软性。这是一种精致高级的趋法，但是具有许多规则的专家系统的“脆性”仍然令人气馁。
- 使得本质“软性”的系统充分复杂，从较高的层次观察时有时将显示出硬性。

我认为，最后一条建议是最令人感兴趣的。它假定了一个不以规则为基础发挥功能的系统，但是在此有某些系统性质可以用规则来描述，如果这些规则被证明是有用的。我弄不明白，为何这种建议不会满足那些喜欢找到结构和模式的人[5]。他们可以不用声称已经发现了基本的组分而作出有趣且非常有用的分类，也不用将基本组分抬高成终极真理(Final Truths)。不过，我怀疑这对于真正的科学家将是一种成问题的见解，在他们看来，如同福多尔和佩利夏因(1984：64)所说，“真理比体面更为重要”[6]。

2.3 联结论与后现代科学

关于科学、后现代理论和复杂性的联结论模型之间的关系，将在最后一章中加以考察。不过，这里需要从认知科学的语境中作出若干评注。

在许多的科学领域，无论是理论的还是应用的，都对分析的、决定论的方法和描述越来越不满意。对这种不安的初期反应之一是，不仅在对实验的诠释上，而且也在对结果的解释上，统计趋法都得到迅速增长。不过，如同模糊逻辑(fuzzy logic)的情况，统计趋法并不意味对于决定论方法的突破。它仍然是一种工具，运用于建立所调研现象的真正机制的过程中。只是在此过程中付出的沉重的代价——把复杂的内部细节平均化——通常被掩盖过去了。

按照关系来进行思考，而不是按照决定论规则来进行思考，对于科学来说并不是新奇的，但过去这一直被看作定性描述的部分，不被看作是定量描述的部分，而自从开普勒(Kepler)坚持主张“进行测量就是进行认识”以来，计算就是必要的。许多现象，特别是在生命科学中、在物理学和数学中，仅仅从决定论的、基于规则的或统计的过程是不可能正确理解的。亚原子过程的量子力学描述本质上是关系的，甚至在更为宏观的水平上，也是关系决定着物质的本性。身体中的碳原子，是完全可以与桌子中的碳原子互换的，它们之间不会有可察觉的区别(Penrose 1989：32)。因此，每一原子的意义并非是由其自身的本性所决定的，而是大量的其自身与其他原子之间的关系的结果。

按照这些例子所显示的，使人感到奇怪的是，对于大脑行使功能的描述明显是一种关系结构，但是仍然牢固地将其附着在原子表征和决定论算法上。其中的缘由之一肯定是，认知科学继承了决定论的、分析传统的方法论框架。我认为，后结构理论有助于我们转变这种立场。

在后结构主义和认知科学之间的相互作用，应使两者都能受益。

在方法论水平上，一种后结构趋法能够支持认知的非算法本性。它可以有助于我们超越那种期望：基于关系的系统模型不那么严谨，而只是有用，尚需找到完全的、决定论的模型。这还有助于说明这种活动的合法性：不要奢望能够填满“大图景”(big picture)，仅是在局部有用。人们还可以认为，在联结论和德里达的语言模型(下一章中将要讨论)之间有一些有趣的相似性。如果后结构的主张有助于使我们得到更好的关于复杂系统的模型，它就会帮助我们消除掉如下的流行见解：后结构概念只限定在理论领域，而没有实际价值。

注 释：

1. 对于特定神经元输出的方程，可以是：

$$\phi = f(\sum_{1}^{n} O_n W_n)$$

式中：

ϕ 神经元的输出

n 与此神经元关联的前面的神经元数目

O_n 第 n 个前面的神经元的输出

W_n 第 n 个前面的神经元相伴的权重

f 神经元的转移函数，几乎总是非线性函数，常常是S形

关于神经网络的详细技术性介绍，参见 Rumelhart and McClelland (1986)。

2. 图灵机概念可能是误导的。它们是抽象的数学实体，提供了描述计算程序的形式系统，计算程序即以离散步骤遵循某种算法的过程，包括执行于冯·诺伊曼型数字计算机上。事实上，图灵机以“读写头”阅读，并以符号打印在(有时是无限长的)“带子”上，这并没有将其“真正”制造出来。关于图灵机及其与计算机的关系的更详细讨论，参见 Haugeland (1985：133～140)或 Penrose (1989：40～86)。

3. 它们无法区别语义网与神经网的分布式表征，也导致了它们使联结论等同于联想论(associationism)(Fodor and Pylyshyn 1988：31)。联结论网络不能对特定思想之间的众多关系进行编码，因为网络中的特定节点并不相应于特定的思想，而在语义网络中则有这种相应。一些混淆是联结论的术语导致的。联结论倾向于将网络中的关系看作是“亚符号”或“微观特征”，好像是有一组微观特征结块在一起，加和成为一个整体特征或符号。当将它与德里达的“痕迹”概念(参见下文)进行比较时，我们能更好地理解何为权重。

4. 早在福多尔等人提议之前，这种研究已经在联结论者中得到了认真的关注，参见例如 Touretzky and Hinton (1985)。

5. 可以用一个类似的论据来捍卫理性的“后现代”形式，这不是基于普遍原理，而是形成于应付——并是关键性的——我们复杂的生活世界中的必然性之中。理性不可能先于意义的生成，而是必须追随它，而意义是复杂关系集的结果。参见 Schrag (1992)从哲学而不是从联结论视角提出的类似论据。

6. 这种在知识和伦理之间的划分将在最后一章中加以讨论。

第 3 章　后结构主义、联结论和复杂性

本章的主要目的，是引入后结构主义以指出它与复杂系统研究的关系。这将通过对由索绪尔(1974)提出的、受到德里达(1976)批评和发展的语言理论的讨论来实现。对于熟悉这种思想主要内容的人，也许希望直接推进到本章的后半部分，在那里对复杂性理论和联结论网络之间的关系进行了讨论。

索绪尔的语言“结构”模型，在复杂系统的研究中是一个里程碑。他的主要见解——意义是通过有差异的系统产生出来的——仍然是一种对复杂系统关系进行概念化的极好方式。他的模型有些“刻板”，但是通过对于关系如何即时相互作用的精致描述[运用了延异(différance)的概念]，这种系统的转换得以实现，从而给我们提供了一种极好的从哲学视角对复杂系统的动力学进行概念化的方式。而且，我希望指出，联结论的归复网络，可以用来为此种一般概念化建立模型。

本书的核心论据因此可以概括如下：

- 复杂性最好被描述为通过大尺度、非线性相互作用而产生出来的特征。
- 由于它是基于关系的系统，对于语言本性的后结构探究有助于我们对复杂系统的相互作用动力学进行理论概括。换言之，在语言中产生出意义的动力学，也可以被用来一般地描述复杂系统的动力学。
- 联结论网络分享着复杂系统的特征，包括由后结构语言理论所描述的那些方面。因此，有可能运用它们（或其他具有类似功能的分布式建模技术）作为复杂系统的一般模型。这些模型是物理上可执行的或计算上可模拟的。

为了对这种论据给出具体说明，我将首先介绍关于语言的后结构概念框架。

3.1 索绪尔的语言模型

大多数的语言模型，特别是乔姆斯基模型，主要关注的是语言的结构。在《普通语言学教程》（Course in General Linguistics）中，索绪尔主要关心的是语言的意义。语词是怎样获得意义的？他的确主张，语言是由一些离散单元组成，他将之称为语符（signs），但不是将特征归结为语符自身，而是关注它们之间的关系[1]。

对于索绪尔来说，语符是由两种组分构成的：能指（signifier）和所指（signified）。能指是语言单元，而所指是它所代表的概念。不过，能指和所指并非是可分离的实体。"语言的语符是一种双面心理实体……其单元并非是某种事物或名称，而是一种概念和一种声形（sound-image）"（Saussure 1974：66）。例如，"树"这个语词在英

语中便是一个能指，而树的概念便是所指。它们共同形成了关于树的语符。语言就是这种语符的系统。

按照索绪尔(1974：67)，语符也具有两种“基本的”特征。第一种最重要：在能指与所指之间并没有自然的连接。树的概念和“树”的语词之间的关系并非是从语言外部赋予的，而纯粹是人为的。这并不意味着，个别的言说者可以为一个概念选定此语词，其关系仅仅是“非促动的”(unmotivated)(69页)，即它的存在只是语言中的一种约定。索绪尔认为，第二种特征是语符在时间中线性地展开。对于他的整个理论，这个特征看起来不那么重要，但是，它的确强调了其所坚持的将语言的言说形式看作是基本的观点，我们稍后将转到这个问题上来。

语符具有意义是如何成为可能的，如果它们的本性是约定的，且同时不是选定的或由个体言说者集体定义的？作为系统的一部分，“有语符之处就有系统”(Culler 1976：91)。语言系统不是由个体言说行动构成的，而是超越了个体使用者的关系系统。这种系统，被索绪尔称作“语言”(*langue*)，作为与投入使用的语言——“言语”(*parole*，Saussure 1974：13，14)相对的概念。因为能指—所指关系是人为的，所以语符并不具有某种自然特性，而是必须从系统的关系中导出其意义。

为了解释这些关系起作用的方式，索绪尔运用了列车的例子，比如说“8点25分日内瓦到巴黎”(108页)。尽管就列车而言，其乘务人员和旅客每天都是不同的，但是“8点25分日内瓦到巴黎”通过其与“8点40分日内瓦到第戎”、“12点整日内瓦到巴黎”或“0点38分孟买到马德拉斯”的关系而保持其特性，在此情况下，无论它是否准确地于8点25分离开，或者它离开后是否到达巴黎这个车站。

列车自身并不具有某种特性；其特性是从关系上确定的。类似地，语言的语符是从它对于其他语符的关系中导出它的意义。能指“棕色”并不是因为它与明确包含着“棕色”本质的概念相一致而具有意义，而是因为它可以与能指“黑色”、“蓝色”、“灰色”、“硬”、“列车”等等区别开来。语符是由它相异于系统中所有其他语符的方式所确定的——“语言中只有差异而无主动项”(120页)。语符是关系网络中的节点。关系不是由语符决定的，相反，语符是相互作用关系的结果。

索绪尔的确承认，能指之间的关系是变化的，因此他区分了语言中的历时性研究和共时性研究。共时性研究将关注系统于时间的某一点，而历时性研究将试图表明系统的一定状态是历史地达到的。不过，变化是高度可控的。索绪尔以两种(似乎是矛盾的)方式强调这一点。一方面，他坚持了语符的不可改变性(immutability)：

> 能指，尽管表面上对于它所表征的思想完全是自由选取的，但对于使用它的语言共同体则是固定的、非自由的。大众在此并无发言权，语言选取的能指并不用另一个来代替。这个看起来包含某种矛盾的事实可以用白话讲：“堆叠起来的牌。”任何个人甚至是希望这样做的人，都不可能以任何方式变更此种已经作出的选取；而且更进一步说，共同体本身也不可能施加如同对于单个词汇那样多的控制；它受到已有语言的束缚。(71页)

语言系统超越了任何个体使用者的选择，因此具有稳定性。因为语言的使用者必须在他们所继承的语言系统中运用语言，而且因为他们对于语言的理解由系统所构成，所以他们不可能与之割裂。系

统以保证其完整性的方式而使得本身长期存在。这就是索绪尔把语言作为“堆叠起来的牌”时所要表达的意思。另一方面，他也坚持语符的可改变性(mutability)：

> 保证了语言的连续性的时间，支配着另一种明显与前者矛盾的影响：语言的语符总在或多或少地迅速变化着。因此，在一定的意义上，我们能同时谈语符的不可改变性和可改变性。归根结底，两种事实是相互依赖的：语符置身于变化之中只是因为它在使得自己长久存在。在所有变化中占据主导的，是旧基质的持续；而不忠实于过去只是相对的。那也就是为什么变化原理以连续原理为基础。(74页)

这种表面上的矛盾，是语符的人为本性的结果。如果语符具有某种自然的或本质的意义，它便会总是保持着完全一样。如果它等同于个体言说者重置其意义，那么当能指变得与另一个能指联系在一起时，意义便会即刻发生变化。这两种可能性都没有反映语言工作的方式。因为关系是人为的，而且只是在系统中才有意义，没有什么“本质”的东西，也没有孤立语符的即刻变化。“不论变化的力量是什么，无论是孤立的还是组合的，它们都总是会形成在能指和能指之间关系的变迁中”(75页)。变化只有在先前的关系与新的一组关系之间产生差异时才可追踪，这是进化过程的结果。索绪尔就语言的进化说了如下的话：

> 没有任何东西比这更复杂。由于它是社会力量和时间两者的产物，没有人可以改变其中的任何东西，另一方面，它的语符的人

为性在理论上承诺了可在语音基质和思想概念之间建立起任何关系的自由。结果是，在语符中结合着的两种要素的任何一个都高度保持它自己的生命，而为他处所罕见；此外，语言在所有可能影响其声音或意义的力量的影响下发生着变化，或不如说发生着进化。进化是不可避免的；没有一种语言可以阻挡这种进化。经历了一定的时期之后，一些明显变迁总是可以记载下来。(76 页)

索绪尔的主张是：语言，甚至是如同世界语那样的人工语言，也只有在它处于不流通的情况下才是可控的。只要它被共同体运用，它就将适应和变化。索绪尔还坚持认为，系统作为一个整体，绝不是被直接地变更的，变化的仅仅是系统的要素(84 页)。然而，变化了的要素将与系统的其余部分相互作用，这就使整个系统最终也会发生变化。这种结果，并非是一个新的、完全不同的系统，而是老系统的一种变换(85 页)。在此也不存在为变化提供方向的外部目的；变化是在语言被运用之处的语境中偶发产生出来的结果。索绪尔(77 页，78 页)相当明白地强调，这种语境是由运用此语言的言说者共同体提供的社会语境。按照一般的复杂系统理论，人们可以说，这些语言系统的动力学，是系统发生着自组织以满足共同体需要的结果(参见第 6 章)。

尽管索绪尔将语言领域描述成其中的要素只是由它们的关系所构成的有差异系统，但他还是坚持了此领域之中的大量的一般性区别。除了能指和所指之间的区别之外，还有历时性和共时性、语形学和音韵学，以及语段学和典范学之间的区别。语段学关系关注的是特定的序列，其中一些语言的语符被用来形成某个句子。典范关系，是被应用于某个语符和所有其他的能够取代它的语符之间的关系。看

来，关系系统并不能满足其描述语言的需要。他不得不加入一系列一般性区别以提供某种更强的结构，甚至尽管他认识到，这些区别没有哪一个能够独自地发挥作用。在此总存在着对立面之间的辩证关系(Culler 1976：50)。

在索绪尔的语言描述中更令人意外的要素是，他坚持认为，言说语言要比书写语言更为基本。对于索绪尔(1974：24～25)，书写不过是语符的一种形象，语符是在言说之中找到其真正的表现的。正是形象使得语符的真正本性模糊到这样的程度，被他称之为“书写的专政”(31页)。如果人们将他最初的关于语符的“心理实体”(66页)的定义加入到这种言说语言具有重要性的主张之中，语言便获得了个人的、主观的本性，而这是与他将其描述为超人格的系统相抵触的。看来，索绪尔一方面否认语符具有本质性本性，另一方面却又坚持认为个体言说者能以某种方式恰好把握它。在下一章中，在考察德里达关于书写的优先性主张时将进一步地对这种张力进行考察。索绪尔坚持认为言说语言的优先性，导致了把优先性给予言说者的心智状态，给予她/他在头脑中的概念，因而也就是给予所指。能指成为所指的记号，所指的表征。不过，随着坚持认为书写的优先性，这种过程就逆转了：现在所指不再具有特殊地位，它成为另一种能指，其意义并非是被运用时存在的，而必须从相互作用着的能指的整体系统中追溯而来。

3.2 仅作为能指系统的语言

索绪尔引入了一种语言系统，其中语言组分不是通过规则来规定其特性，而是从它们与所有其他组分的关系中导出其意义。初看起来，这明显地发生了某种相当根本的转变。然而，如果人们考虑到

索绪尔既坚持系统的稳定性又坚持系统在线性时间维度上的进化性，那么事情就清楚了，即系统的机制可以被给予某种相当传统的描述。哈兰(Harland 1987：136)以如下方式将它描述成“同步性系统”：

> ……所有的语词稳定地静处于一个整体同步的系统中。同步性，即系统所取得的平衡，来自于语词在同一时间的相互推挤；总体性，即系统所取得的平衡，来自于系统中不存在语词滑落所需的空间(即内部空隙)，也不存在语词分散所需的空间(即周围空隙)。对于完美的平衡态，语词需要在一个封闭的空间中被致密地捆绑起来。

从根本上讲，索绪尔把语言理解成系统，其中每个词都有其位置，并相应地有其意义。此系统的确在进化，但是保持在某个近平衡态。不过，复杂系统，如语言，并不运行于近平衡态，系统中组分之间的关系是非线性的和动力学的。语词，或者语符，并无固定的位置。语符之间的关系，并没有稳定到每一语符都是可以被精确地决定的。只有语符之间具有某种“空间”，相互作用才是可能的。可能性，总是要大于可实现性(Luhmann 1985：25)。语符的意义，是在语符之间的空间之中“游戏”的结果。复杂系统中的语符，总是具有意义的富裕，即只有一部分潜在的意义在特定的情形下得到了实现。

我们如何来谈论具有这种特征的系统呢？如果“同步性系统”的组分并非固定的，那么我们如何来描述其相互作用的动力学呢？要试图回答这些问题，先关注德里达对于索绪尔的解构也许是有帮助的。

德里达对索绪尔关于语符描述的批判，联系着他对西方哲学整个传统的趋势的批判，他将其称为“在场的形而上学”(metaphysics of

presence)。在索绪尔的情形，在场的形而上学是由他坚持语符具有两种组分(即能指和所指)所确认的，其中所指是心智的或心理的(Saussure 1974：66)。这意味着，语符的意义对于使用它的言说者是呈现的，从而对意义由差异系统构成提出了挑战。这也就是为什么索绪尔坚持言说的优先性。语言一旦被书写下来，主语与它的词汇之间就出现了距离，使得意义变得不确定了。

不过，德里达觉得索绪尔没有理由担忧这种事态。他坚持认为，主语与词汇之间在任何情况下总是存在着距离；语符的意义总是不确定的，甚至在我们言说之时也如此。因此，所指(或"心智"组分)绝不会有任何即刻的自呈意义(self-present meaning)。语符就只是语符，它从其他语符导出其意义。这样的观点导致了在一定意义上，语符剥离了它的"所指"组分。由于所指通过关系系统构成，它的功能也正如能指。一个所指不是别的，就是另一个能指，它必须在能指之间的无尽的相互作用中取得其位置。意义绝非仅仅是呈现，因此我们不可能逃离解释过程，甚至当言说者立于我们面前时也是如此。这也就是为什么德里达选择书写——那种总是已经具有了距离的"能指的能指"(Derrida 1976：7)——来作为语言系统的模型。

> 现在从人们把已决定的语符、言说以及强音(fortiori)、书写这四者的总体性当作固定不变的组分这一时刻起，人们必须排除在能指或能指秩序之中的任何自然从属关系和任何自然等级。如果"书写"能指描述，特别是一个语符的持续组成(这是书写概念中唯一不可归结的核心成分)，书写一般包含着整个的语言语符领域。(Derrida 1976：44)

语符的解构，即其心智成分的去除，与主体和意识的解构密切相联。在两种情况下，解构都在抵制在场的概念。如果不存在其内容是直接地呈现给意识的能指，而仅有其意义是不稳定的、富裕的能指，那么，意识的内容就变得过剩了；它不可能是完整的。主体不再处于可得以呈现的意义的控制之下，其自身就由能指的运作而构成。

以系统理论的语言，索绪尔仍然将语言理解成某种封闭系统，而德里达希望把语言看作是某种开放系统。否认在场形而上学时，“在内”和“在外”之间的区分也是有问题的。意义在语言中生成，语言之外的场合是不存在的。只要有意义，就一定有语言。我们不可能将语言与其描述的世界分隔开来。“外在承载着内在的关系，即一般而言只不过是外在性。在外的意义总是呈现于在内，无论是否将在外限定于在外，反之亦然。”（Derrida 1976：35）只有当取消了在内和在外之间的区别时，系统才能成为开放系统。

如果语言系统是如同德里达所建议的开放系统，如果关系总是以某种不可预见的方式发生着变化，那么，哪里可能言说这些关系呢？在一定的意义上，我们不可能说它们有任何的永恒和特殊，这也一般地适用于语言。然而能指的运作导致了“稳定之囊”（Stofberg 1988：224），否则交流就无从开始。在这些囊中，对于关系进行更严密的分析是可能的，只要理解到稳定并非是永恒和彻底的，意义始终是能指之间的相互作用的结果。德里达在对这种相互作用的一种解释中使用了两个概念：“痕迹”（trace）和“延异”（différance）。他强调，它们实际上既不是概念也不是语词（Derrida 1982：7），它们不可能被赋予完全的意义。也许德里达是在将不必要的东西复合起来，但他的意图看来是阻止这两个术语获得固定的意义。正是由于它们的不稳定性，使得我们对于语言能进行更一般的言说。我将试

图以有意义的方式来描述这两个术语。

索绪尔从语符与系统中所有其他语符的关系来规定语符的意义。例如，语符“棕色”的意义不仅是由它相异于语符“蓝色”、“绿色”和“红色”来决定的，而且也是由它相异于诸如“狗”、“弹簧”和“拙笨”来决定的。语符不具有仅仅属于它自身的组分；它不过是所有贯穿于它的其他语符的痕迹的集合。这是被德里达强调的。语符是一种没有任何积极内容的实体。因为它只不过是由关系所组成，所以它仅仅是由痕迹所组成。其所含的意义是明确的，组成了一个特定语符的痕迹，不是从其他语符散发出来的，因为它们是自足的，并从而能给出具有某种积极的内容。恰恰相反，所有的语符都是由差异系统所组成的。

> 差异的运行包含着综合，并阻止了在任何时刻或以任何方式，一个简单的要素只是呈现于本身或自身之中，而且仅涉及自身。无论是书写的还是言说的话语，都不存在这样的要素：在其与其他要素（其自身也不仅仅是呈现）没有联系的情况下而作为语符发挥作用。这种联系意味着，每一“要素”——音素或字形——都是参照其他序列或系统的要素中的痕迹而构成的。这种联系，这种交织，就是文本，文本只是通过其他文本的变换而来。无论是要素还是系统，其中都没有任何东西是可以简单看作在场或缺席。无论何处，存在的都只是差异和痕迹的痕迹。（Derrida 1981：26）

痕迹的来源没有某种固定的参考点，无论从时间上看还是空间上看都如此。共时地或历时地追溯痕迹的起源，都是不可能的。人们可以试图追踪到痕迹的种种途径，但是绝不可能到达某种还没有被差

异所分化了的起点或起源[2]。

将语符描述成仅仅由痕迹所组成是问题的一方面，但是就痕迹的相互作用的“机制”作一些言说也是可能的。在此，我们对于复杂系统的动力学的理解，能够通过德里达的“延异”概念得到促进。这是一个复杂的概念，其中有几层意义。首先，它指的是将语言描述为一个有差异(differences)的系统。痕迹是差异的痕迹。在有差异的运行中，意义产生了出来。然而，因为这种运行总是处在不停息之中，意义绝不会是最终的产物，而是不断地被延迟(deferred)。对于一个语符，只要一定的意义产生出来，就会在系统之中得到反响。通过种种的回路以及途径，痕迹中的这种扰动就会反射回到该语符自身，从而改变了它的“最初”意义，即便这种改变是难以察觉的。因为痕迹不会处于优先的位置，“延迟”(defer)一词不仅可以从时间意义上阅读，也可以从“顺从”的意义上阅读。每一痕迹都不仅仅是延迟了，而且也是顺从于任一其他痕迹的。

痕迹之间的相互作用的机制超出了差异和延迟。此机制并非是系统的积极特征，也不是积极作用的结果。延异过程滞留于主动和被动之间的某处。语符是由系统产生的，但同时也涉及系统的产生过程。系统的特征是作为痕迹延异的结果而涌现出来的，而非是系统的特定组分的本质特征的结果[3]。

一个并非从其多义性中出现的延异要素，是“空间”的要素。为了使语符通过痕迹和延异而相互作用，正如我们上面已经见到的，它们不可能被紧紧地堆叠起来。行动作用需要有“空间”。主动与被动之间的滞留在这里也是明显的：只有有了一定的空间，并在延异动力学中加以保持，延异才可能发生。因此，延异既具有空间特征也具有时间特征(参见 Derrida 1982：13)。

痕迹和延异两个概念被用来讨论语言的内部工作状态，或者说——由于语言成为任何相互作用着的语符系统的模型——被用来讨论所有复杂系统的内部工作状态。在随后的内容中，我将探讨后结构理论和复杂性的网络模型之间的特殊联系。

3.3 网络和系统

大约在索绪尔建立他的语言理论的同一时期(大约在19世纪与20世纪之交)，弗洛伊德提出了一个关于大脑功能的模型。这部早期的著作被称为《科学心理学方案》(Project for a Scientific Psychology)，是在作者身后才发表的(Freud 1950)，不如他的成熟著作那么有名，但对于完整地理解弗洛伊德却是必不可少的。而且，只要涉及大脑功能问题，它大体上与当代的神经学理论是相容的(参见Pribram and Gill 1976)。

弗洛伊德模型是由神经元构成的，神经元通过大脑中的传递能量的渠道而相互作用。这种能量既来自于身体的外部(感知)，也有内部的来源。渠道抵制着能量流动，除非它被经常地使用(在此我们得到了“使用规则”或“赫布规则”的最初表述)。大脑的特征是由其中能量流的种种模式所决定的。这种模型有两个方面是值得注意的。第一个方面，记忆的作用应该被强调。“记忆”在此是指大脑的物理条件：哪种途径被突破(“被推动”)，以及哪种没有。记忆并非是由意识主体所施行的某种认知功能，而是大脑(作为身体的一部分器官)的无意识特征。记忆是为大脑的所有功能建立条件的基质。

弗洛伊德模型的第二个重要特征，涉及神经元的作用。神经元并不由自身而获得意义。记忆并不驻留在神经元之中，而是驻留在神经元之间的关系之中。弗洛伊德(1950：300)宣称，这种关系是一

种差异。因此，我们得到的是一个结构上等价于索绪尔语言模型的模型：一个差异系统[4]。把德里达对于弗洛伊德和索绪尔的诠释作为一条线索，我们可以利用对于语言的后结构趋法这种理论等价物，建立一种关于相互作用神经元的网络动力学的描述。

德里达利用“痕迹”概念指出了语言系统中的每一组分对于所有其他组分的影响。痕迹概念密切地联系着记忆概念——如上面所描述的物质的、非主观意义上的记忆。在神经网络中，记忆功能由神经元之间的关系权重来执行。因为这些关系的“分布式”本性，一个特定的权重并不具备构思能力，而只是在大量的相互作用的模式中才获得了意义。因此，“权重”和“痕迹”这两种概念在这种语境下可以互相描述的假设将是富有成果的。从神经网络把权重看作痕迹（在德里达的意义上），有助于我们理解网络中有意义的模式是如何主要从权重的条件中形成的。而从语言把痕迹看作权重，则有助于我们将它们设想成某种并非短暂易逝而是尽管稀少但现实的实际东西。

类似地，德里达的延异概念也能够用来描述复杂神经系统的动力学。这种类推如下所述。如果某个神经元的集合（无论是真正的还是人工的）产生出某种活动模式，该种活动的痕迹在网络中产生反响。当网络中有回路时，这些痕迹经过了一定的传播滞后（延迟）之后会反射回来，变更了（导致差异）在最初形成的活动性。由于复杂系统总是包含回路和反馈，滞后的自变更将是网络的特征之一。这种特征与延异概念有很大的共同性——差异表明的是差别和延迟，滞留于主动与被动模式之间，具有空间和时间两种成分（Derrida 1982：1～27）。按照关于痕迹和延异的后结构“逻辑”，语言中（或大脑中神经元）的语词不可能凭其自身而具有意义。意义由系统的组分之间的动力学关系所决定。同样地，神经网络中节点不可能凭其自身而

具有意义，这即是分布式表征概念的核心含义。意义是从涉及许多单元的活动模式中产生出来的，模式是在大量权重的动力学相互作用中形成的。

转向现在实际运用中的用于解决问题的神经网络(主要是在模式识别领域中)，有必要作出某些限定。实际的网络一般是设计来执行特定任务的。它们的神经元的数目有限，通常相互关联的模式也有限。权重值以及神经元的转移函数也是有限的。而且，网络是为特定问题而设计的，权重通常也只是在学习阶段中变更的。一般说来，它们不够灵活，不可能用来解决大范围的问题。在此意义上，神经网络是结构的而非后结构的，能够相当合适地以索绪尔的术语来加以描述。然而，后结构概念的确变得重要了，我们是不是应该希望将现在的网络的能力加以扩展，特别是在关于 AI 方面。网络若要在模仿人的行为方面更为成功，将不得不具有更大的灵活性。它们应该能够在新奇条件下进行创新，即它们将必须能够超越先定的限制。要实现这样的目标，从后结构主义获得的见解可以作出很大的贡献。

后面将对复杂系统的自组织和信息表征进行更详细的分析。在此之前，作为一段插曲，我将扼要地考察一位学者——塞尔——的一些论点，他看不到后结构趋法的有用之处。

注　释：

1. 在这整个介绍中，“语符”这个概念可以看作在功能上等价于联结论的网络中的“单元”或“节点”。

2. Culler (1983：96)提供了如下的例子：

> 如果一个洞穴人由使用某种咕哝声来标志“食物”从而成功地开创了语言，我们就必须假定，这种咕哝声已经与他种咕哝声相区别了，于是世界已经被划分成了“食物”和“非食物”的范畴。

3. 他自己对于延异概念的简洁概括，参见 Derrida (1982：89)。

4. 有关更详细的比较，参见 Cilliers (1990)。

第 4 章　塞尔迷雾

在继续扩展一种关于复杂性的后现代视角之前，我将考察一位明确地拒斥这种趋法的思想家的论点。塞尔也许并不单纯是一位分析哲学家，从他认为复杂系统（如语言和大脑）遵循了严格的基于规则的趋法来看，这将变得很清楚。他相信，从这种框架中，有可能一方面排除德里达，另一方面推翻人工智能纲领。

1980 年，塞尔在权威的《行为科学与脑科学》(Behavioral and Brain Sciences)杂志上发表了一篇文章，叫做“心智、大脑和程序”(Minds，Brains，and Programs)。其目的在于提供一种摧毁性论据，用以对抗那种假定一台运行程序的计算机甚至被说成能够思维的观点，该文章激起了一场自从那时起就经久不衰的争论。由于此杂志采取“开放的同行评议”，所以在发表这篇见解独特的文章的同时还有 28 篇批评文章，后来的各期中刊登的文章更多，而塞尔作出的所有答复都是信心十足的(Searle 1982，1985)。这场争论不仅在其他的杂志（P. M. Churchland and P. S. Churchland 1990；Denning

1990；Harnad 1989；Rapaport 1986）和书籍（Bierwisch 1990；Münch 1990；Penrose 1989）中继续着，而且原文被选进了许多关于认知科学和人工智能的重要文选中（Haugeland 1981；Hofstadter and Dennett 1982）。英国广播公司（BBC）在播出了一系列演讲之后，辑集出版了《心智、大脑和科学》（Minds，Brains and Science）（Searle 1984）一书，这场争论弥散到了普及性出版物中，甚至是定期地出现。尽管有无数的反驳，塞尔却保持了一种平静的信心十足的态度（Searle 1990）。

为何此文受到了如此之多的关注？智能计算机——或人工智能（AI）问题——必定会引起广泛的兴趣。除了以牢固的和良好的科学研究程序为支持之外，它抓住了公众的注意力主要归功于那若干本出色的普及性著作（Hofstadter 1980；Penrose 1989——前者是“支持”，而后者是“反对”）。AI 又是引起广泛哲学兴趣的问题。它的失败或成功，对于我们理解复杂系统（包括人类）以及它们是如何工作的具有重要的启发意义。它也适合于列入到一种一般的哲学传统和策略中，这种传统和策略以“分析的”一般性名称而著称。若对 AI 问题获得了某种结论性的答案，则大量理论的状态，特别是涉及大脑和心智的都会受到根本性影响。所以，塞尔的文章讨论了一个众说纷纭而又激动人心的问题。

但是，这样的一篇文章，在一个发表物已经多得难以应付的领域，为何会引起如此之多的兴趣呢？所有的原因都是错误的。塞尔并没有提供某种特别明确的论据，而随后的争论也并没有真正深入地阐明此问题。所有希望取消 AI 的人都盲目地接受了它，而它又受到所有 AI 支持者的激烈反对。论战是如此轻巧，这引起了我的怀疑。持续 15 年的指责和对峙的传奇，具有典型的仇敌世家的味道。在我对于他的见解的分析中，我将试图表明，塞尔并非是人工智能的主要

反对者，而恰恰是这个阵营的成员。

4.1 中文房间

塞尔否认任何强形式 AI 的论据，采取了思想实验(Gedanken-experiment)的形式(Searle 1980：417～418)。他要我们想像，一位讲英语而完全不懂中文的人，处于一房间中。给此人一堆中文符号以及规则说明书，说明书中正式地描述了中文语法。对于他来说，这些符号是一堆毫无意义的胡乱书写，但是通过规则，他可以操作它们。如果从外面给他一组符号，他可以生成一组新的符号并送出来。从房间的外面来看，这位房间中的人似乎是能够讲中文的，因为他对提示和问题作出了恰当的反应，而房间中的人并非像他理解英文那样理解中文单词。他不过是遵循了一组形式规则。因为这正是计算机所做的——遵从一组规则——但任何计算机都不能“理解”自己在干什么，因此不能认为计算机会思维。

这里勾画的实验梗概，对于 AI 共同体而言是相当熟悉的。它非常类似于计算机先驱阿兰·图灵(Alan Turing 1950)提出的关于机器是否具有智能的测试。从根本上讲，该测试由与隐藏于终端的计算机进行的扩展对话而构成。如果在对话的这一端，你不知道是在与一个人还是与一台计算机对话，那么就认为该计算机通过了图灵测试。尽管赋予语言这种突出重要性受到了某些人的质疑，如哈纳德(Harnad 1989)就宁愿引入一个称作全图灵测试的机器程序，但是图灵测试仍然保持了被普遍接受为机器智能的基准。塞尔主张，已经确凿无疑地表明，一台计算机可以通过理解中文的图灵测试，但它却仍然不可能真正地理解或思考。这就意味着，这个结果使得图灵测试作为适当的智能测试是不可信的(Searle 1980：419)。

正是在这一点上，塞尔似乎是提呈了一个清晰而简明的论证，可能为有趣的讨论提供一个出发点。然而，他立即遭到了一系列批评，他反驳了这些批评以捍卫自己的观点。这把他引入了困境，当我们把其中的一些回答以及塞尔的反驳联系起来考察时，这一点就变得清楚了：

- **关于系统的回答**(419 页)：说一个人进行了操作但不能说他理解了中文，如同说一台计算机的中心处理器不能理解。这忽视了最为重要的规则手册(或程序)。整体的系统(包括人、规则手册和符号)可以被认为是理解中文的，因此，一台计算机加上其程序也是可以理解的。对此，塞尔的回答是主张那里没有任何东西阻止此人记忆规则手册。然后他就相当于是整个系统，却仍然无法理解遵循着规则自己在干什么。
- **关于机器人的回答**(420 页)：假定中文程序形成了控制某个机器人系统的一部分，此机器人不仅给出了适当的回答，而且执行了适当的行为。这样的机器人便会有真正的理解。塞尔回答说，控制程序并不能对机器人通过种种“感知器”获得的输入进行区分。对于此程序，所有信息都只是由形式地操作的符号组成，而这并不意味着任何理解。
- **关于大脑模拟器的回答**(420～421 页)：假设对大脑工作原理(可以讲中文)模拟到其中的功能过程都得到了反映这样的程度。这样的模拟必须能够理解中文，否则便会被迫否认原先讲中文的、其大脑被模拟的、能够理解自己所讲的东西的人本身。塞尔对此仍然是无所畏惧。他指出，这种模拟能够用一套由管道和闸门构成的精致系统来实施，其中管道和闸门代表了大脑

结构，直至神经元和突触的层次。他继续指出：

> 那么，这种系统中的理解力何在？他用中文作为输入，模拟着中文大脑的突触形式结构，并将中文作为输出。但是此人肯定不理解中文，水管也不理解，如果我们试探着接受我所设想的荒谬观点，把人和水管莫名其妙地关联起来从而能够理解的话，请记住，在原则上人可能使得水管的形式结构内在化，并在他的想像中承担着所有的“神经元激发”。大脑模拟器遇上的问题是，它是在模拟关于大脑的**错误东西**。只要它仅仅模拟的是神经元在突触激发的序列的形式结构，它便不会模拟什么大脑的东西，即其因果性质，**它产生意向状态的能力**。而形式性质对于因果性是不充分的，正如水管的例子所表明的：我们可以得到所有的从有关神经生物学因果性质中剥离出来的形式性质。(421 页；楷体为引者所强调)

对塞尔文章的反响中，有大量其他的批评，但是他利用这 3 个回答(关于系统的回答、关于机器人的回答和关于大脑模拟器的回答)来作为他回应那种种批评的基础。大多数的相反论据都被归结到这 3 个问题类型之一，从而他以对那种类型的回答提供的预制解答来进行反驳。这种轻而易举在于：塞尔的主张是同义反复的，因此它们能够用来对抗任何形形色色的诘难，或支持任何与强 AI 抗衡的论据。在塞尔著作的结尾处，这种同义反复变得明显了。他对自己论据的概括，是通过追问大脑所具有的东西即使之能够理解和思维的东西是什么？它们如何不同于程序？塞尔的回答是，它们具有意向性：“某些大脑过程对于意向性是充分的。计算机程序的初始化本身对

于意向性是不充分的条件”(417 页)。人们也许会问，何为“意向性”？在“中文房间”的文章中，他并没有提供定义，而是在他的取名《意向性》(Intentionality)(Searle 1983)的著作中才附带地以大写的方式标出了这个词，人们可以读到：“意向性是心智状态和事件的性质，关于世界的事态都借助于它或受到引导，或受到围绕，或具有对象和状态。”(1 页)。他的如下表述是有启迪作用的：

> 意向性完全不是某种本体论问题。例如，信念究竟是什么？……
>
> 信念是某种心理学模式的命题内容，其模式决定了心智对于世界的满足感取向，其命题内容决定着一组关于满足的条件。如果我们不打算失去关于它们的内在意向性的见解，意向性状态必须以意向性术语为特征。但是如果问题是“信念的存在模式和其他的意向性状态是什么？”那么关于这个世界是如何运行的，我们目前所知道的一切也就是：意向性状态既是由大脑的结构引起的，也是由大脑的结构所实现的。(14 页，15 页)

“如果我们不打算失去关于它们的内在意向性的见解，意向性状态必须以意向性术语为特征。”这种主张，要么是同义反复，要么是毫无意义。我无法领会的是，在何种意义上，一个计算机程序在一定条件下必须要打开比如说闸门时，比起大脑中的信念来，“关于”世界中的某种东西要少一些，除非大脑的基质有某种神秘的东西。这样的假设，看来好像由“中文房间”论文中的如下陈述所证实：“石头、盥洗室使用的纸、风和水管(所有的材料都可能用来构建计算装置)都不是一开始便具有意向性的材质，而只有某些具有与大脑

同样因果能力的东西，才可能具有意向性”（Searle 1980：423）。如果你剥掉形而上学，那么你就将终止于总体的循环论证之中：大脑不同于程序是因为大脑具有意向性而程序没有。何为意向性呢？意向性是使得大脑不同于程序的东西。

循环和矛盾此后还在继续。在“中文房间”论文的结尾处，我们发现了如下的段落：“‘机器能够思维吗？’我自己的观点是，只有机器能够思维，而且的确是只有非常特殊种类的机器即大脑以及具有与大脑同样因果能力的机器才能够思维。”（424 页）这些“因果能力”是什么，它们从何而来？塞尔并没有给出任何线索，而是继续写道：“大脑当然是一种数字计算机。由于万物都是数字计算机，所以大脑也是。”（424 页）在提出了“关于大脑模拟器的回答”（420 页）之后，无论如何精确地对大脑进行模拟，模拟仍然是一种形式系统，并因此不可能具有意向性！这就够了。从所有这些之中获得的最合理的结论看来是，只有数字计算机才能够理解中文，而任何学习过中文的人则不可能理解中文[1]。

那么，为什么这种瞎话却导致了持续近 20 年的活跃争论呢？我认为，因为塞尔是 AI 兄弟会的内室成员。他采用的是他们非常熟悉的专业术语的论据，使得他们乐于向它射击。这种情况是这样轻而易举地发生了，因为他们都共享着一套基本的前提。这些前提，也就是将被仔细加以审查的东西。

4.2 中文房间的框架

在对塞尔隐含的假设进行分析之前，我希望先考察一下这种看法，即他的论据具有形而上学的味道。罗蒂（Richard Rorty）在讨论塞尔的论据时，将整个问题带入了一种有趣的视野之中（Searle

1980：445～446）。他认为，塞尔在处理其论据时，采取的方式与原教旨主义的天主教为捍卫圣餐变体论采取的方式完全相同：

> 假定除去神话色彩的蒂利希神学家坚持要求，我们对圣餐的思考不是从物质变化而是从信徒的生命意义方面进行的，正统的捍卫者将回答道："自然的与超自然的区别不可能只在旁观者的眼中，而必定是内在的；否则所有的旁观者都有能力将事物看作超自然的。"（445 页）

正如正统的神学家所预知的，什么将自然与超自然相区别，什么是超自然的"因果能力"。塞尔预知的是，计算过程不可能处于认知状态，因为它不可能具有同样的"因果能力"。

在回答罗蒂时，塞尔（453 页）把此种类比较斥为"总体无关"，因为罗蒂仅仅是提供了以相似为形式的论据。而论据的真理性取决于前提的真理性，于是塞尔立即陈述了他自己所采用的前提："它们是这样的东西，如同人们具有诸如信念、欲望和目睹经历这样的精神状态，而且也如同人们有大脑，他们的精神状态因果地是其大脑运行的产物。"（453 页）我假定，假如这些前提是作为"常识"，那么不会有人从内心怀疑它们，因为塞尔并不为它们进行任何辩护。停留在他的论据的框架之中，人们便会认为，只要把这些前提和几条新鲜的意向性混合起来就能通过一种琐碎的办法得到塞尔的所有结论，但也许这也正是质疑框架本身的时刻。这些"常识"前提——无论我们是否采纳它们——便遮盖了其他的前提。这些都没有明确地叙述过，但是它们对于整个的思想实验却是中心问题。它们涉及语言的本性。

塞尔的整个论据所关注的"中文房间"目标，是包含着对符号进

行操作的规则手册。尽管的确有某些评论者谈到了规则手册的可能大小和复杂性，但是看来对于这种手册的可能性则没有太多关心。塞尔完全没有担心其实用性，信口开河地将其当作已经存在的东西。这种说法中至少有三个重要假设，它们暗含于塞尔的语言理论中。

首先，他假定了可以为自然语言建构起形式语法，并以查阅表的形式提呈出来。第二个假设与第一个假设密切有关，预设了在语法和语义之间的明确分离。塞尔对此说得很清楚，“规则说明了按照其句法而非语义对符号进行纯粹形式的操作”（Searle 1984：32）。塞尔的“规则”概念，与乔姆斯基在早期表述中即在生成规则（Chomsky 1957）意义上的使用方式，显然是一致的。

塞尔的第三个假设涉及意义的本性。在《言语行为》（Speech Acts）中，他拒绝“意义就是使用”的观点，认为它过于模糊（Searle 1969：146）。取而代之的是，“意义即是意向”（Searle 1983：160）。我们已经遇上了他的意向概念的模糊性。它是如何支持关于意义的理论的？塞尔的主张是：“意向性正是精神状态的特征，人或其他[原文如此]，这使得这些状态来代表其他事物”（Searle 1979：197）。塞尔支持一种关于意义的表征理论。这在他的早期著作中业已清楚（参见 Searle 1969：第 4 章），而且他仍然这样认为：“精神状态是‘符号的’，至少在它们代表世界上事务的对象和状态的意义上是如此。”（Searle 1985：743）[2]

我们于是看到的是关于语言和心智两种概念的混合体。无论是塞尔或是我都不会对此有异议。我们都同意，这些问题是纠缠交织在一起的。不过，塞尔应该对此涉及的复杂性更为敏感才是。在讨论大脑和计算机能够做什么而不能做什么时，将语言活动对回答的问题约简成简单的行动，“以 squoggle-squoggle 取代 squiggle-squiggle”

(以一种形式的胡乱涂鸦取代另一种形式的胡乱涂鸦)(Searle 1984：32)，是非常俗套的主张。在我看来，在如下的问题上也是有矛盾的，他指责对手持一种二元论(Searle 1980：423)，但自己又认为在语法和语义之间有明确的区别，并在被指出之后仍然坚持语义是大脑的特有特点，即意向性的结果[3]。

最尖锐的对立来自同一派别中的重量级人物之间。因此，塞尔为何介入了这场特殊的战斗，反响为何如此炽热，也就应该清楚了。他和他的对手都有这样的共同信念，即持有关于形式的基于规则的语言理论的信念，其中语言的结构能够以纯粹语法来描述，意义是表征的结果。他们不持有同样的心智理论反倒是没有道理的。他的对手非常明白这一点，但是他顽固地拒绝放弃意向性概念。在他自己看来，只要他拒绝赋予其形式描述的可能性，意向性就将保持为形而上学的概念。如果你(以及我)不接受这种语言的基本特点，整个的思想实验就无法进行下去。

让我清晰地叙述一下我关于“中文房间”论据的看法：除了其内在矛盾和盲点之外，整个论据是不切题的，因为它建立在忽视了语言复杂性的语言理论的基础之上。塞尔无动于衷地从哲学层面上将复杂系统与简单系统无差别地等同起来。对于他来说，“以 squoggle-squoggle 取代 squiggle-squiggle”包含着操作语言所必需的基本原理。余下的只是细节问题。我认为，任何从这种简单理论中抽取的结论，对于大脑或机器的问题都谈不出更多的东西，是无助的。

在考察了塞尔的“中文房间”之后，我们对他轻视后结构主义就不会感到惊讶了。

4.3 塞尔对后结构主义的拒斥

由于塞尔感觉到自己应是言行理论(speech-act theory)的重量级人物，针对德里达对奥斯汀(Austin)进行的解构式阅读，他就会责无旁贷地责问德里达。当然，塞尔了解奥斯汀真正的意思，因此能够纠正德里达的错误。两者之间展开的争论，再一次显示了塞尔有过于简单化的倾向。

奥斯汀(1980)在他建立言行理论过程中的贡献之一是，主张言词的意义取决于言说行动实施所处的语境，这里的“语境”包括适合于行动的社会习俗。不过，有必要找到某种使得言说行动得以成功的措施或运气。一般接受的解答是建议，行动的成功是由言说者的意向与给定的情形或语境的协调一致。

在一篇取名“署名事件的语境”(Signature Event Context)的重要论文中，德里达的确是抓住了这一点[4]。在他看来，言说者的意向和语境都没有充分固定不变，从而难以决定任何形式的交流的正确含义。一旦它构成了“记号”——无论这些记号是言说者发出的声音还是书写下来的文字，任何表达也就脱离了其最初的羁绊，因此大致宽松地说，意向不可能“伴随”着语言去控制其进行的解释。交流的行为，总是一种“分发”(Derrida 1988：2)。另一方面，语境也不是客观地给定的。德里达试图表明，“为何语境绝非是可以完全决定的，或者说，为何其决定绝不可能是完全的、充分的”(3 页)。

语符可以重现(即人们在不同的环境中一再使用它)这一结构必然性，是对于语符可以在发送者和接受者都完全缺席的情况下进行操作的保证(7 页)。德里达将其称为语符的“可重现性”(iterability)。这也是为何德里达主张，讨论语言的基本本性时更适当的比喻是“书写”而非“言说”的部分原因。言说者的突破是意义的“起源”，预

定的接受者的缺失排除了交流过程之中的“在场的连续变更”。为保持言说行为的正当，发送者、接受者和消息绝不可能互相“在场”。当我们解释一条消息时，我们只能进行——解释。要使一个语符成为可解释的，就必须能辨识出那个特定的语符，而不是其他的语符。不过，其意义的再现绝非是一成不变的。语符具有其历史，历史影响着其意义。每当语符被使用，便会与语言系统中的其他成员相互作用，因此其意义就变迁了，有时难以察觉，有时却又明显。可重现性概念既描述了语符的稳定性也描述了其意义的动力学，是与痕迹和延异概念密切关联的。

德里达对语言的分析，对于我们讨论复杂性具有重要意义：它导致了“归根结底把编码当作终极的规则系统的权威地位的终结”（8页）。总之，它反对关于基于规则的语言描述的可能性。

塞尔(1977)对德里达的论文作出了反应，认为德里达完全误述了奥斯汀的见解，他混淆了可重现性与一般援引性，使言说行为仍然是“在特定的历史语境中的可确定年代的异常事件”（208页）。言说者和听言者可以理解新事物，是因为他们是“被我们称作语言规则的规则集合的主人”（208页）。

德里达对此已给出了一个长长的回答，不仅为自己的见解辩护，而且首先是对其进行证明。他表明，尽管塞尔作出了最大的努力，但也不可能阻止他的语词的意义发生变更。这只要引用(即重现)塞尔这些语词因语境的不确定而发生变化，便是证明。塞尔声称正确地解释了奥斯汀的著作，在文章的开端处又作了一个版权声明，德里达将这两者联系起来，指出塞尔试图运作限于少数人的店铺——因此是一家有限公司。他还在文中通过引证塞尔的反驳论文中所使用的每个词（包括版权声明），嘲笑了关于版权的声明。通过这一系列

的错误陈述，德里达看来是暗示他们之间从来不曾有过真正的交流，因此他们之间的对立也就不曾发生。

德里达与塞尔的争论，特别对于“中文房间”的论证，以及更一般地对于我们讨论的复杂系统，都具有重要的含义。首先，它否认“意向性”能成为语言的意义或意识的基本特征终极的参照点。德里达并不否认意向性的存在：“意向范畴将不会消失；它将具有自己的位置，但这个位置不足以统摄言说的整个场景和系统”（Derrida 1988：18）。在塞尔那里明显具有形而上学维度的意向性概念，因此不可能用作对于复杂系统的涌现性质的现成解释。此讨论从而确认了塞尔对于复杂性缺乏敏感性，这也是德里达完全意识到的，“人们不要为了进行复合的愉快而将事物复合起来，”他认为，“但是人们也不要在不存在简单性之处进行简化或假装确信这种简单性。如果事物都是简单的，那么语词便早已传播遍布”（119 页）。

“中文房间”对于我们的深入思考具有某种影响吗？我认为，影响不多。继续攻击或捍卫塞尔的论证，将不会使任何人真正受益。因此中文房间，实在是过分不现实。而且，基于生成规则的形式模型来趋进复杂系统，会涉及一种根本性困难。机器是否能够思维，对此我们并没有一锤定音的论据。这取决于我们在未来对于复杂系统建模取得成功的程度——这是一个评判的标准仍然缺乏的条件性问题(contingent matter)。然而，我的确倾向于认为，仅仅以基于生成规则系统的机器将能够处理例如自然语言的可能性是不存在的。当然，本书的观点是，联结论网络在这方面有着一定的前景。帕特里夏·丘奇兰德(Patricia Churchland)和保罗·丘奇兰德(Paul Churchland)夫妇(1990)感到，这种网络的威力是充分大的，足以克服塞尔的反驳(假定其仍然需要改进)，但是塞尔(1990)认为联结论没有任何新东西，这对于他是

一种必然的结论，因为他持有一种反对大脑模拟的老论据。

联结论的大多数支持者，包括丘奇兰德夫妇，都在 AI 研究的目前的主流理论的语境中来捍卫自己的观点。其中最重要的是意义的表征理论。至少对丘奇兰德夫妇来说，神经网络仍然以传统方式表征信息。而我并不认为，联结论模型目前的形式是解决我们面临的建模困境的万能办法。还有大量的工作要做。然而，我的确认为，在做工作的过程中，从后结构主义理论可以获得富有意义的洞察。

我将最后概括一下对于塞尔论据以及他的表述方式的反驳意见[5]。这不仅仅是过于轻易地使得意向性在你的所有论据中最为重要，而且当追问它是什么时，你认为完全可以说那不过就是"关于"大脑。同样轻而易举的是，将语言归结为"以 squoggle-squoggle 取代 squiggle-squiggle"。复杂问题需要复杂描述，还要一定的谦逊。塞尔不承认任何的智力恩惠，也缺乏历史眼光。在一本冠以《意向性》书名的著作中，他只是顺便提及胡塞尔一次(Husserl)(Searle 1983：65)。思想概念，总是以好像它们都是从他的脑门中完美无缺地蹦出来的形式表示。他对于批评的反驳中，也不时体现着这种自负。当前面所讨论的德里达的两篇论文被收入选集中时(Derrida 1988)，塞尔拒绝自己的文章被选编入其中。他允诺进行的反驳最近出现在一卷批评性论文集中(Burkhard 1990)，编者所收到的所有文章中都是"'这位名人'的简短'Hm'"(26 页)。对于德里达(Derrida 1988：29)的问题"塞尔在哪里？我知道他吗？"我们也许只能回答道：这有任何意义吗？

注　释：

1. 塞尔的论据导致尚克(Roger Schank)得出结论，只有从书写程序的人(即 AI 研究人员)中才能找到能够理解的人(Searle 1980：447)。

2. 更详尽的讨论，参见 Marletti (1990)。

3. 这个涉及谁是或谁不是二元论者的问题，可能形成不同的独立探索路线。霍夫施塔特(见 Searle 1980：433)对于塞尔在这方面状况的评价是相当清楚的。他认为，意向性只不过是塞尔用作"灵魂"的词汇。

4. 这篇文章以及德里达答复塞尔的批评的文章，还有扩展了的"后记"，其编辑者的署名为有限公司(Limited Inc.)(Derrida 1998)。我参考的是这个编辑版本。

5. 在已经写完了这一章后，我见到了 Globus (1991)写的一篇文章，也提出了一些异议，其中一些与我的意见是类似的。

第 5 章　表征的一些问题

复杂系统的模型必须如同系统本身一样复杂。它们也必须仿效这些系统的编码和记忆以适合于其环境以及其应付环境的能力。我们已经提出，经典表征理论并没有对这种过程提供适当的描述。表征带来了什么样的问题，为何它们对于理论和实践如此重要？对于我们的复杂系统模型，不同的表征方法的含义是什么，它们如何影响这些模型的设计和运行？这些问题构成了本章的基本内容。

一个表征理论，本质上是一个意义理论。它试图通过定义我们语言的词汇或大脑的结构与世界之间的关系，来解释这些词汇或结构是如何变得有意义的。然而，这种理论也必须能够解释这些关系是如何产生的。我们可以以数学为例。数学符号本身并没有什么意义。它们通过定义(例如，令 x 代表水的温度)而获得了意义。于是就说，该符号"表征"了它所代表的东西。在一个明确定义的问题中，符号与其表征的东西之间的关系可以明确地加以说明，从而容易解释其结果。例如，给定牛顿定律，仔细地运用正确的表征，经典力学的问题

便可以高精度地获解。如果一系列的计算提供了答案 $t = 10.9$，你为了解释此结果只要搞清楚 t 代表着什么(这是需要预先设定的)。

不幸的是，当我们处理复杂问题时，其清晰的定义和明确的边界不易找到，使用符号来表征某事物也就不再容易了。从传统的数学观点看，这些问题会被看成是难以处理的，但是计算机能力的显著提高使得关注的目标回到了诸如自然语言、人的感官能力(例如模式识别、计算机视觉)的建模问题、高级认知过程的模拟(推理、决策)和精致的机体运动行为(机器人学)等复杂问题。在这些情形中，表征的过程远非是直截了当的。语言中的一个词是如何代表某种事物的？专家是如何提出他们的知识的？模式包含着某些明确的能被模式识别机表示出来的特征吗？这些问题是与如下的事实混合在一起的：许多的这种系统，特别是生命系统，如果没有非常严格的形而上学的含义便不可能赋予外部设计者某种重要的角色。如果造出某些超验的原理(如上帝或宇宙观念)，这样的表征理论的首要前提便会乞求于过多的问题。因此，有必要对符号和世界之间的关系得以建立起来的过程加以解释。这些是希望建立复杂系统的计算模型的人们所面临的问题，而它们并非是全新的。它们是对于语言、心智和世界关系进行哲学反思的长期传统中的一部分。

在计算建模者的语境中，看起来好像这个表征问题能够以大致两种方式进行探询。第一种方式，我们有经典的计算机科学和 AI 的基于规则的趋法。在此，逻辑演算的适当性是广泛接受的，是一种理性主义趋法。在第二种方式中，我们有联结论或神经网络趋法，这更多地具有经验的色彩。对于这两种趋法，表征问题都是一个中心问题(Brooks 1991：139；Guenthner *et al*. 1986：39)。在神经网络的情形下，表征问题被翻译成了如下的问题：网络结构(“结构”

通常指的是权重的值和关联模式)及其所引起的现象之间的关系是什么? 网络如何表征它所建模型的范围的信息? 是否有可能对成功的网络结构进行分析以确定它是怎样解决此问题的? 我们是如何运用某种先验知识来预先确定网络的结构? 能获得这样的信息吗? 基于规则的表征和网络表征可以互相归结吗?

这所有的问题，对于神经网络的设计者和使用者都具有实践意义。 本章对它们的探讨，试图澄清那些最为重要的决定着任何特定模型的框架的预先数学考虑。 这将通过批判经典表征概念的方式而进行。 随后将对联结论网络中发现的分布式表征概念进行分析。 也许有人会认为，这种趋法使得整个的表征概念都面临挑战。 如果能够有说服力地表明联结论并不依赖于强表征理论，这就将进一步强调后结构理论对于我们理解复杂系统的重要性。 因此在对表征问题的技术性讨论之后，我们将对一些后现代哲学家[特别是鲍德里亚(Baudrillard)和德里达]处理它的方式进行考察。

5.1 真表征

只有一种语言适用于表征信息——无论是陈述上或是程序上——这就是一阶谓词逻辑。只有一种智能方式来处理信息——这就是运用演绎推理方法。

科瓦尔斯基(Kowalski)(引自 Guenthner *et al*. 1986; 41)的这段话，浓缩了计算机科学对于表征问题的态度。 也许人们应该承认作出如此大胆的陈述是勇气可嘉，但是这种对于逻辑的盲从则是简单化了。 而且，尽管有一些限定和保留，大多数研究人员都沉浸在这种复杂系统的建模方式中，这或明或暗地支持了科瓦尔斯基的信条。

我并不试图贬低逻辑的重要性，但认识到逻辑的局限性也很重要。我将试图通过分析一种大体上接受这种信条的表征理论，把这些局限显现出来。这种批判，将有希望为联结论趋法提供支持——这种趋法并没有采用演绎推理和一阶逻辑作为其出发点。

5.1.1 功能主义

两种复杂系统——人脑和自然语言——是特别有意义的。因为它们看来是解决(或绕过)了表征问题，它们对于解的可能性提供了存在的证明。如果这两种系统得以成功地建模，此问题便大体上得到了解决。而且，这两种系统看来是解不开地纠缠在一起的——其中一个模型的成功也就意味着另外一个的成功。那也就是为什么大多数心智模型都是基于——或结合着——语言理论的。我们所要进行的讨论也不例外。

在计算建模界，主流的心智理论被称为功能主义。普特南(Hilary Putnam)是功能主义的奠基者之一，尽管他后来不再持这种信念。

> 我也许是第一位写下计算机就是心智的合适模型的论文的哲学家。我将这种信条的形式取名为"功能主义"，在这一名称下它已经成为当代心智哲学中的一种主导性观点——一些人称其为正统观点。(Putnam 1988：xi)

心智功能的计算机比喻是这样的：一台工作中的计算机有两种组分，机器本身(硬件)和在此机器上运行的程序(软件)。软件(运行于一个不同的逻辑层次上)可以从一个硬件转到另一个硬件，并仍然以完全相同的方式运行。心智与大脑之间的关系，被认为与此类似。

大脑构成了“硬件”，可以运行不同的程序。这些程序产生出“心智”和效果。正如计算机软件并不限制于在某台特定的机器上运行那样，心智同样也可以运行于其他的硬件上，只要它们具有适当的计算能力即可。这最后一点需要加以说明。

按照功能主义，机器(大脑)的一定的物理状态实现一定的功能状态。这向我们提供了同一事件的两种独立的描述：一个在物理水平上，一个在心智水平上(Sterelny 1990：5)。心智在大脑中运行，但原则上也可以在其他的具有适当计算能力的机器上运行。在功能主义看来，数字计算机，或更准确地说图灵机，便具有这种适当的计算能力。

在设定了功能主义的基本前提之后，理论上提出了两个问题。第一个涉及物理系统施行功能状态的适当性。一桶水是否也够格？它是充分复杂的，足以表现无数状态。为了去掉此种不够格的系统，斯特瑞尼(Sterelny 1990：10)提出了关于适当的物理系统的两个标准：

- 它们是设计出来的。
- 它们具有目的论解释。

斯特瑞尼选取进化作为设计者而避开了第一条标准，但是忘记了进化也能产生网状联系和无数的其他“无心智”系统。我认为，此种陈述只能被看作是具有表面价值：接受功能主义理论意味着存在独立的外部设计者。我在后面还将返回到这一点。第二条标准也是有问题的。斯特瑞尼并没有真正地解释其含义是什么，只是说诸如一桶水或太阳系这样的东西没有目的，而大脑和身体则具有目的。我觉得

他在此忽略掉的概念是“意向性”，即与塞尔(1980)在挽救其“中文房间”论证中所使用的概念是相同的。你无法说它是什么，但是当你发现它时却可以认得它。两个标准都具有某种形而上学的味道，这应该为功能主义的支持者所承认。

功能主义提出的第二个问题，涉及同一物理事件的两种独立的描述之间的关系。是什么赋予了一定的物理状态以一定的意义？这个问题将在下一节中讨论。

5.1.2 思维语言

物理系统的功能状态，只有在它们代表着某种事物时，才被赋予了意义。没有表征，功能主义就是毫无意义的。语言的语法结构表征着语义内容，大脑的神经学状态表征着一定的心智状态。在两个例子中，系统的状态和概念意义之间联系的建立都要归结到表征。

> 没有表征，就不可能有信息敏感性。没有表征，就不可能有对世界作出反应的灵活性和适应性。为了从世界中学习，并将所学以新方式运用于行动，我们都必须要能够表征这个世界、表征我们的目标和选择。而且，我们必须要从这些表征中进行适当的推理。(Sterelny 1990：21)

表征是把两层次的描述——符号及其意义——联系起来的过程。这是如何进行的以及不同的表征是如何相互作用的，都是表征理论要清楚说明的。尽管关于它有一些变形体，但一个被普遍接受的表征理论，其核心之处已经被福多尔(1975，1981)以乔姆斯基(1957，1972)的语言学理论及其心理学含义为基础提出来了。我将回到其中的一些变形体，但是首先我们必须按福多尔的方式对表征的基本特征

进行一个概括。

福多尔模型的基本主张是，我们使用一定的内心语言——通常被作为“心思”（mentalese）（Fodor 1975）——进行思维。心思并非等价于我们所讲的语言（例如法语或德语），而是先于语言的。我们的心思能力并非是我们获得或学习的某种东西，而是我们大脑的一种天生的能力。如同其他的语言一样，心思是独立于中介的。所以，我们的头脑中没有书写的或言说的词汇；语言施行于我们的神经结构中。

思维与语言共享着两种重要的特征。首先，语言是生成的——任何句子的长度都可以增加而不毁坏句子的意义。其次，语言是有结构的——它有语法允许我们对先前无关的语言要素进行推理。如果 A 和 B 共享着某种特定的对于 C 的关系，语言允许我们在无需经验确认的情形下推导 A 和 B 的关系。类似地，思维可以无限地联结起来，可以用来进行系统地推理。以此出发，福多尔认为，心思必定与语言共享着可给予生成性和结构以最好解释的东西：形式语法[1]。这个结论被斯特瑞尼（Sterelny 1990：26）以如下方式加以概括：

> 因为这个模型，以及任何以它为基础的模型，都需要一个代理来将世界表征为如其所是以及也许所是，并从那样的表征中得出适当的推理。福多尔主张，行为者必须具有似语言符号系统，她能够表征关于环境的无限多和无限复杂的实际的和可能的状态。没有适当的表征手段（即思维的语言），她就不具备这种能力。

关于思维语言假说的这种大致概括，如果我们考察一下由卡明斯（Cummins）（参见 Cummins 1991； Morris 1991， 1992）提出的更形式化的明确说明，对它的认识就将更深入一些。这同样也是“基本上

是以经典 AI 为模板”(Morris 1991：2)的认知的计算机模型。卡明斯将此模型概括成“塔桥图”(参见图 5.1)。图中，下面的空心箭头表明了执行层，上面的空心箭头表明了对该层的解释。各符号的含义则是：*i* 和 *o* 分别是物理系统的输入和输出状态；*g* 表征着相应于某个因果规律的函数，它将执行层的 *i* 转变成 *o*；*I* 是解释函数，由此 *i* 和 *o* 被赋予了“内容”；*f* 是将输入内容转变成输出内容的函数。按照莫里斯(Morris 1991：3，4)所言，当且仅当如下条件成立时，此模型才起作用：

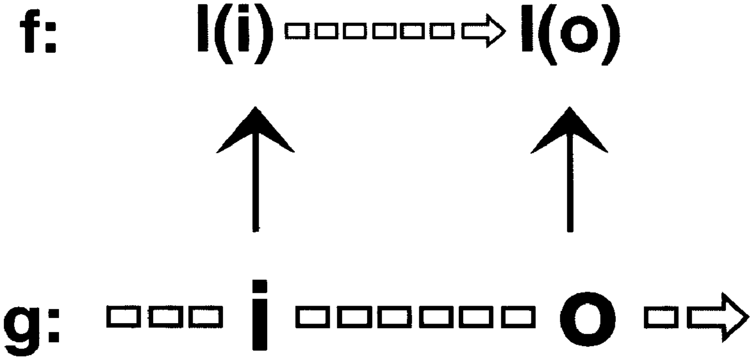


图 5.1　塔桥图

（i）输入和输出具有内容。这个条件规定了在此必须有表征。系统的状态 *i* 和 *o* 必须代表着某种东西。

（ii）在解释层具有可理解性，使得输入内容转变成输出内容。这个条件说明了 *f* 的本性。通过规定 *f* 必须具有“规则”，将会在“心智”层上被接受的此解释被限定为可被讲清楚的基于规则的解释。莫里斯(1991：4)认为，这是基本的条件，因为“除非我们阐明了某种使得从一种内容转变成另一种内容的规则或函数，否则我们就没有资格说获得了分离的单个过程”。由于顶层是不能不要的，所以，没有表征模型就不能工作。

（iii）在执行层上，输入和输出在非语义上是可区分的（没有参照内容），以及

（iv）存在着非语义的因果规律，解释了系统中那些输出是如何从那些输入中产生出来的。

条件（iv）说明了 g——某个因果规律——的本性，并与条件（iii）一起确认了执行（语法）和解释（语义）层的完全独立性。这意味着，I 不可能是物理系统本身的结果。

此模型的两个进一步的特征必须被提到。第一位的是，此模型是同步性地工作的。它保证了一个非历史的表征概念（Morris 1991：6）。第二位是，语义层在逻辑上是先于句法层的。构成了“适当的”语义解释的东西的定义（Morris 1991：7）如下：

$$g(x)=y，当且仅当 f[I(x)]=I(y)。$$

按此定义，执行是对于在语义层发生的真实过程的模拟。我觉得，这两个特征尤其揭露了该理论中嵌入的形而上学内核，不仅是一种笛卡儿二元论，而且是一种抽象的非历史的观念论。莫里斯（1991：17～19）认为，这种阐述恰当地描述了福多尔的表征理论。

这概括了经验表征理论的基本架构。人们也许还会加上这种看法：由于此种理论主张符号具有抽象的、非偶发的意义，因此是彻底的现代主义理论。我对它的评价和批评，将出现在关于联结论模型的讨论中。经典的表征理论，对于从计算上执行的形式系统（即可以按照谓词逻辑完全描述的系统），为我们提供了一个良好建构的、连贯的模型，但它对于描述诸如自然语言和人脑这样的复杂系统则是不适当的。在转到联结论模型之前，我将考察取自其早期代表人物之

一的普特南对于经典表征模型的一些异议。

5.1.3 普特南对经典表征的批评

在《表征和实在》(Representation and Reality)一书中，普特南改变了他原先的功能主义主张。他认为最主要的问题是表征的作用：

> ……乔姆斯基理论化的关键思想是：(1)普遍语言的思想，(2)先天性假设，(3)新近的模块性思想。于是，语义层次上乔姆斯基理论将采取的形式就相对清楚了（而福多尔的理论确实采取了这种形式），即使其细节可以有种种变形。一个语义层的乔姆斯基理论将认为，存在着心智和大脑中的“语义表征”；这些是先天的和普遍的；所有我们的概念都可以分解成为这种语义表征。这就是我希望摧毁的理论。(5 页)

福多尔和乔姆斯基的表征模型，被普特南称作最新形式的“心思主义”(mentalism)。此理论的吸引人之处在于，它简便地将信念欲望心理学（或民族心理学）与计算建模区别开来。他认为，心思主义“正是思想史中为更一般趋势所采取的最新形式，这是将概念看作科学上可描述的心智或大脑中（‘心理学上真实的’）实体的趋势”(7 页)。普特南认为，这样的趋势完全是误导的。

尽管他的确并没有真正提供可替换的建议[如 Goodman (1991)指出的]，但是普特南的确得到了 3 个反对心思主义的强论据，对此我希望加上第 4 个。

(i) **意义是整体的。** 反对心思主义的表征的第一条论据认为，在符号与其指称的事物之间不仅有一种关系，而且有多种关系。

意义是由演奏团的活动产生出来的，而不是由某种独奏者产生出来的。“整体”一词在此并不承载着大量的负担（它对于有机生长的紫花苜蓿的萌芽完全是两码事）；它只不过是指向这样的一种事实，即我们在处理意义时面临着关系之网。这种整体论趋法必须要反对两种理论倾向。

首先，整体论是反对实证主义的。实证主义的倾向是希望把意义还原成“认识论上更原始的”语言所阐述的定义（Putnam 1988：8），其组分能够以明确的方式（以感性的或至少是以可观察的角度）被赋予意义。其次，整体论反对借定义而把术语的定义固定下来的倾向，因为“大多数术语都是不可能被定义的——或至少，如果定义是指永久地固定某种绝对把握术语意义的东西，那么术语是不可能被定义的”（9 页）。

从整体论的论证来反对表征，首先不是因为它是某种抽象，而是因为它是一种过于简单化。它表达了一种空想的乐观主义，即认为我们最终将使得事情到位，而无需在复杂斗争和偶发相互作用过程中亲自动手。

（ii） **意义是规范性概念的一部分。**随此而来的论据并非是后结构的——即认为意义是由复杂关系（这还将包括力量的关系）所产生并歪曲的。普特南的主张不过是认为，因为术语的意义不可能按照某种基本的或物理主义的概念——这已经降至最低标准——来定义，那么总是会涉及信念的因素。我们大多数人都没有见过电子，无论是直接的还是间接的，甚至那些声称获得了其存在证据的人也必须将他们的主张置于某种理论框架之中。我们对于电子是什么的理解，是以某种共享信念系统为基础的，而且因为并非我们所有的人都具有同等程度

的关于电子的知识，因此这些信念是有差异的。

假如一个系统中的符号只有一种正确而完全的意义，如果遗漏了那种本质，那么交流就会被破坏。在通常情形，成功的交流并不必盯在本质上，而毋宁是相互给予对方不确定的益处。这通常被称作“宽容原理”(principle of charity)。所有的相互作用都依赖于宽容，因为“当我们进行解释时，总是必须至少在信念的某些差异上持不确定态度”(13 页)。

(iii) **意义依赖于环境**。这个论据主要是指向福多尔的先天性假设和观点，即认为我们的环境产生出了在3万年的进化之前无法预见的概念(15 页)。而且，我们在意义的建立过程中发挥作用，改变着有差异的环境之间的复杂相互作用，会影响意义的产生，而无论有何种的先天性假设。我们不是历经着我们世界的连贯的、抽象的主体；我们是由我们的世界产生出来的，我们改变着我们的世界，也被我们的世界所改变。

(iv) **意义是一种历史概念**。尽管普特南并没有把意义的时间方面作为一个独立的论据，但它实际上成为所有这上述三方面论据的一部分。在此，它再次提示，复杂系统不仅仅是由外部环境所决定的，而且也是由系统的历史所决定的。这些方面以非线性的归复方式相互作用，而归复方式是以一阶谓词逻辑而非形象延伸来加以描述的。

普特南并没有真正表明如何能够得以克服这些反对意见，但是这些反对意见必须认真对待。显然，“经典的”表征对于解释复杂系统如何贮存和利用信息并非是合适的模型。如果另外有一个替代方案，它就必须考虑到上面讨论的意义的四种特征。接下去要讨论的

是，某种联结论的解释在这方面也许更行得通。

5.2 联结论

这里不打算重复基本的联结论理论（参见第2章中的介绍），但让我们简要地概括一下神经网络的最重要的特征，再转向神经网络是如何“表征”信息的。

5.2.1 神经网络的特征

神经网络由大量相互作用着的节点或“神经元”的集合体构成。每一个神经元都接受许多其他神经元的输入。所有的联结都具有某种“强度”，这被称作联结的“权重”。这些权重具有实数值，可以为正（激发）、为负（抑制）或为零（意味着相应的两个神经元并没有发生联结）。神经元自身都是非常简单的计算单元。它们计算着自身具有权重的输入，并通过某种非线性转移函数来传递出这种值。这种函数通常呈S形。

网络中的一些神经元作为输入单元接受从外部来的信息。类似地，一些神经元作为输出单元，在此可发现网络计算的结果。在简单的网络结构中——如多层感知机或后向传播网络——神经元是层状排列的。于是可以有输入层，一些中间层（通常称作“隐含层”）和输出层。如果向网络提供一个输入，那么通过网络的过滤，就将产生出一个输出。由于神经元本身都是类似的，所以由网络所实施的变换都是由权重值决定的。

这些权重是如何决定的？从根本上讲，网络是通过向其展示范例而得到训练以执行某些任务的。如果网络必须执行分类任务，例如将树木分类，那就赋予它那些可能与树木混淆的树木和事物的样本。在学习阶段，每次提示（这采取了数值阵列的形式，如在视觉膜

上的每一坐标的强度，或在屏幕或照片上的每一坐标的像元值)，都有相应于该输入的正确的输出值。网络于是自动地调节其权重值，来减少在输入与输出之间的距离。这些提示是连续进行的，直到网络收敛到一组权重值，使得网络可以将树木和非树木的种种样本区别开来。如果训练的样本是合适的，网络还将能够对它以前没有遇到过的树木样本进行概括分类。

实际中最常应用的神经网络——采用后向传播算法进行训练的多层感知机——是一个简单的网络，具有上述的所有基本特征。

5.2.2 神经网络中数据的贮存和操作

在上面的讨论中已明确的是，任何两种网络之间的差异将主要由权重值决定[2]。任何特定网络的特征都是这些值的分布的结果。这些权重如何对从外部世界获得的信息进行编码，一方面取决于权重数值之间的关系，另一方面也取决于提呈给该网络的信息的类型和数量。这是一个重要的问题，我将通过考察三种网络来详加讨论。

第一种要考虑的网络，其中节点代表着特定的概念，而权重代表着它们之间的关系。一个简单的例子是，使得特定节点代表一群人的成员。如果两个人是有关的，它们之间的联结为正；如果无关，那么为负。于是我们得到了一个自动关系识别器。激活任何节点，所有的代表着相关人的节点都会被激活。这种网络在传统 AI 中是相当著名的，被称作语义网络(semantic network)。节点和联结都有特定的解释，而且由于它们“代表着”某种事物，因此就运用了传统的局域表征理论。

第二种要考虑的网络，其中排除了在一些节点和特定概念之间的必然关系。相当简单的模式识别网络(如用来识别树木的多层感知机)，就是这种网络的一个例子。这种网络将在隐含层中有大量的神

经元——至少达到输入层的数量级，每个神经元都与所有其他输入节点关联着。输入网络的信息因此分布于所有的隐含层神经元，使得任何一个神经元都不代表任何特定的对象。这也导致了这样的事实，解决一定问题所必需的隐含神经元的数量可能非常大。如果一个网络中有 120 个隐含神经元，其性能与 100 个神经元的网络相当，则这并不意味着该问题的 20 个特定的方面没有被较小的网络所考虑到，而只是意味着，信息现在分布在少于 20 个单元的范围中，可能伴随有网络性能的略微降低。

在这种网络中，节点与权重都没有特定的解释。编码的信息并非是连贯概念的表征。表征通常都被认为是“潜符号”(sub-symbolic)或编码“微特征”(如 Rumelhart and McClelland 1986)。任何权重其本身都是没有任何意义的。而且，每次参加进来的所有权重，都处于某种模式中。这个编码过程被称作“分布式表征”。分布式表征的特征和重要性，及其联系后现代理论的方式，在本章后面还要更详细地加以讨论，但现在有必要先引入上面提及的第三种网络。

对于第三种网络，可接着前面来考虑多层感知机，这里将隐含神经元数目减少到仍然能够识别所有训练模式的绝对最小值。这种网络可以发展出两个有趣的特征。第一个特征是，当观察其权重值的分布时，明显的是权重倾向于要么是高度正值，要么是高度负值。关于第二个特征，当提呈了任何的输入模式，隐含节点都倾向于要么是猛然打开，要么是猛然关闭。在这样的网络(我将其称为“最小网络”)中，隐含节点开始对输入模式的特定特征进行编码；他们开始代表某种特定的东西。由于权重要么是强烈地激活(“是”)，要么是强烈地抑制(“非”)，它们开始对表征着特定概念的节点的逻辑关系进行编码。含义是清楚的：尽管我们这里使用的是以正常方式训

练的神经网络，但我们通过除去网络中的冗余性，强迫使之采取了一种局域性表征。

在这些不同种类的网络之间划定清晰的边界当然是不可能的，这是由于分布性将取决于问题的复杂性。对于一个问题，50个隐含节点也许可以执行高水平的分布性了，而对于一个更复杂的问题，这也许仅仅构成了一个最小网络。在此，重要的是要把握住，当运用一个小的网络，仅仅能对单元之间的要么接通要么关闭的逻辑关系进行编码时，是谈不上分布式表征的。另一方面，当网络中神经元过分多时，还有一个严重的问题。当一个网络过分威力强大，它会过分地适应训练数据，从而概括能力就很差了。当我们考察这些思想的实际含义时将讨论这个问题，但这里我们首先将给出更贴切的分布式表征的特征。

5.2.3 分布式表征

在这一节，我希望处理直接与分布性有关的神经网络的三个特征，以及对于分布式表征的三种批评。这些将有助于澄清我们对此概念的理解。

第一个重要的特征是，分布式网络固有的处理复杂性的方式。在运用网络来解决某个复杂性问题的过程中，并不必明确搞清楚此问题的结构。在传统的计算解题中，为了构建某个可编程的算法，这种理论是明确需要的。因此，在所有其他活动之前需要先有理论建构。然而，当处理复杂系统时，要构建起这样的理论，即便不是不可能，也将是极其困难的，因为这里涉及大量的以复杂的非线性方式相互作用着的因素。在这种情形下理论的建构涉及将复杂性大规模地约简，而这具有导致不适当系统模型的抽象的高度风险。由于神经网络以隐含方式对大量以非线性方式分布的因素之间的关系进行编

码，对完全的、明确陈述的理论的需要也就消失了。当一个系统充分大时，它将具有充分的冗余，来保证将所有这些因素同时地进行编码。在实际情形中，问题被归结成以适当方式将信息提呈给网络(这方面将在下一节中给予关注)。

这导致了极其重要的一点：用作复杂系统模型的网络，必将与系统本身有同样的复杂度。因此，构建起网络在干什么的理论，也就如同构建起系统在干什么的理论一样困难。如果有某种能够辅助系统行为描述的一般原理，那么就该有可能通过对于网络行为的分析而发现它们。不过，如果系统是真正复杂的，同等复杂性的网络也就是此种系统的最简适当模型，这意味着对它的分析正好与对系统本身的分析同样困难。这对于从事复杂系统工作的科学家来说具有严肃的方法论含义。一个复杂性化简了的模型，也许容易运行，甚至也许提供了一些对于系统的更经济的描述，但是其相应的代价是必须认真考虑的。

分布式网络的第二个特征，涉及它们对解进行概括的能力。在已经进行特定任务的训练之后，它们还应该能够处理相关的却不同于训练样本的新输入。由于新的输入将具有训练样本的部分而非全部特征，因此训练样本不可能以运用网络中特定节点的固定不变或完全受限的表征来编码。在真正的分布式表征中，所有的节点，以及其相关的权重，都参与所有样本。这意味着，无论何时，最大的信息量都可以在所有的节点中获得。因为局域式表征只有在抽象之后才成为可能，无论这种抽象是某种先验理论还是最小网络的受限工作的结果，因此这种表征进行概括的能力都将随着抽象水平的提高而降低。如果任何时候可以考虑的仅仅是若干明晰的特征，那么任何不共享这些特征的样本都不可能被考虑到。为了对复杂样本进行有效

的分类——在此有大量的微特征必须加以考虑——某种适当的分布式表征是一项前提。

分布式表征的第三个重要的特点，涉及这种趋法的鲁棒性。当某个特定的特征概念由特定的神经元进行编码，那种概念或特征的丢失便会对特定神经元造成损害。不过，若表征是分布式的，便不会有网络的任何特定特征与任何特定的神经元捆绑在一起。即使神经元受到了损害，也只会对网络的总体运行造成轻微的影响。网络越是分布性的，则降低程度越轻[3]。对于一个处理理想化抽象的系统，鲁棒性没有那么重要，但是对于必须应对真实世界偶发性的系统，它就是至关重要的。

联结论受到了许多的批评，特别是来自民族心理学和经典表征支持者的批评。这些批评常常是有缺陷的，特别是由于分布式表征概念要么被忽视要么被误解。福多尔和佩利夏因(1988)运用语义之网来作为神经网络的例子，克拉潘(Clapin，1991)将神经网络等同于经典冯·诺伊曼(von Neumann)机器中的CPU！在接下去的段落中，我将讨论一些批评，它们至少对于分布式表征的本性是敏感的。这些批评取自斯特瑞尼(1990：187～189)，他捍卫的是一种“思维语言”类型的表征理论。如果转变一下理论视角，许多此种反对分布式表征的批评实际上都是指向了其优点[4]。

第一个批评涉及权重的地位。在斯特瑞尼看来，“分布式表征与附加其上的微特征之间的关系是很成问题的”(188页)。当我们从子概念的群簇来定义一个概念时，例如，当我们使用一簇概念{水果，酸的，黄色的，……}来定义“柠檬”——“我们即返回到了概念组成分析的理论，……一种令人绝望的、悬而未决的理论”(188页)。与此相联系，斯特瑞尼还有关于权重功能的问题：

功能的和非功能关联之间的区别是无法界定的，甚至在原则上也是如此。分布式网络中两个节点之间的正连接，也许指的是：某种构成性连接（例如，关于老虎的网络中的类似猫的）；某种通常的连接（在此种网络中的食肉动物）；某种意外的、诱发的连接（在此网络中的亚洲人）；或只不过是相关性的连接（在我的例子中，一个喜欢黑色和橘黄色的特殊球队）。（189 页）

显然，“微特征”概念在此成为一种误导。权重设定了给予构思性内容以特定解释。原则上，斯特瑞尼仍然采用的是语义网络。在一个真正分布式表征中，任何特定的权重的信息内容都是这种分散过程的结果，不可能说构成了任何特定事物。显然不存在不同类别（通常的、构成的、诱发的，如此等等）的权重。

把分布式表征的本性看作仅仅是运用组分簇来实现的概念或特征，也是一种误解。那只不过会是对于局域表征系统的改进而已。你现在拥有的并非是一组具有特定语义解释的符号，而是另一种也许有更多符号的解释。分布式表征允许你放弃这种整体理论框架。你不是在以概念或概念簇来工作，而是使用在子概念层次上的关系系统来工作，这子概念层次不可能是给定的，或换句话说，是不需要语义解释的。我将返回到第二种批评。

第二种批评涉及表征的连贯性和稳定性。“分布式表征并没有给予你某种不变的、独立于语境的表征”（187 页）。由于两个系统的起始状态可以不同，它们也就绝不可能准确地代表相同的概念。如果你将不变的、与语境无涉的表征看作是必要或是应该存在的话，这当然是一个问题。关于意义的后现代或后结构理论，通过主张共时性和历时性语境不仅是必要的，而且是意义的构成因素，来否认这种

主张。因为对于送进网络的信息量并不存在理论的极限(假定网络具有足够的容量),因此可以包含尽可能多的内容。而且,因为信息是分布的,在概念和语境之间不必作出明晰的区分,它们是一起被编码的;语境总是表征的一部分。

关于分布式表征的最后一条批评促使我们分析其限度。斯特瑞尼认为:

> 是否所有联结论系统的表征都是分布式表征,这仍然是不清楚的。……假定节点对于节点的影响是局域的,假定不存在将节点群看作整体的处理器,那么网络中的分布式表征便相当于某个外来者对于此系统的观察。(188 页)

这些使斯特瑞尼感到焦虑——"我希望被告知,为何我应该将分布式表征看作所有系统的状态"(188 页)——而我们却可以轻松应对之。考虑到分布式表征的种种特征,我们现在已清楚了,表征这个概念正遭到破坏。分布式表征不是传统意义上的词汇意义上的表征。它分布于表征系统的所有组成之中。在此没有"代表"着某种事物的符号,也没有它们之间的语法关系,系统本身也完全不需要语义的解释层。我们既不需要某种中心处理器,也不需要外部的观察者。我将在本章最后一节返回到这种结论的哲学重要性上,这里先简单考察一下发展联结论模型的两种努力,它们都没有放弃表征概念。

5.2.4 拯救表征的两种努力

联结论已经对认知科学产生了充分的影响,它引起了从直率的厌恶(Fodor and Pylyshyn 1988)到极度的欢迎(P. M. Churchland 1989)

反差极大的种种反响。在这一节中，我希望考察两种对于联结论持赞成态度的反响。第一种由劳埃德(1989)提出，试图将联结论与传统的表征理论协调起来；第二种由丘奇兰德夫妇提出，发现联结论与消除论(eliminativism)之间有一种天然联盟。

劳埃德选择理解心智所必须解决的表征问题作为中心问题。他希望建立一种表征理论，这种理论既是分析的("对部分的理解先于对整体的理解"，Lloyd 1989：8)，也是还原的("一种解释表征是如何以非表征的部分构成的理论")(11 页)。这种理论与我们理解的联结论已经是不一致的了。而且，劳埃德看来对分布式表征的本质的理解有局限性。当他描述作为其主张的框架的元理论时(11～22 页)，这一点就很明显了。他的元理论由一些可运用于任何表征理论的约束所构成。我希望对此说，这些约束使得他的解释强烈地倾向于传统的表征理论，并掩盖了分布式表征的更根本的含义。让我们扼要地考察一下这些约束：

(i) 表征必须是精确的。它必须有可能鉴明被表征的对象(13 页)。

(ii) 表征必定是专注的，即它具有专一性。它必须要剔除一些对象并绕过另一些对象(13～15 页)。

(iii) 表征必定是"可连接的"。富有意义的表征是由独立的有意义的部分构成的。"总之，原子性和可连接性交织在一起；表征的系统是凝块，最小的凝块可以接连起来形成较大的凝块"(16～17 页)。

(iv) 表征是不对称的。一棵树的画代表着该树，而该树却并不代表这幅画(17 页)。

(v) 表征必须起着认知的作用，即它必须能够满足意向功能（17 页）。

(vi) 表征必须是通过进化过程而形成，即是能够通过自然选择而发展的过程（18 页）。

(vii) 表征必须可以通过还原性理论来解释（19 页）。

除了约束条件（iv）对于此论据是中性的，以及（v）看来是承载了要比仅仅是表征的特征更多的东西，其余的条件都是与真正的分布式表征不相容的。因为他对所有表征的定义都是使用分析的、原子论的术语，所以劳埃德的约束如同上一节所讨论的斯特瑞尼的批评，阻碍他明白分布式表征根本不是真正的表征，并阻碍他明白这是一个好消息。

另一方面，丘奇兰德夫妇对语义水平上的消除主义没有提出什么问题。在他们看来，联结论模型是某种天赐之物。然而在分布式表征仍然形成表征的意义上，它们确实仍归属于表征。他们认为，网络构成了以"一种组织化的关于种种原型可察觉情形的内部表征的'图书馆'"（P. M. Churchland 1989：207），或是在高维空间运行的"灵活查阅表"（P. S. Churchland and Sejnowksi 1992：138）。更进一步，内部层次上的神经元被看作是"复杂特征的探测器"（P. M. Churchland 1989：123）。尽管这是关于神经网络运用于明确定义的问题时对隐含层作用的流行解释，却掩盖了对于前一节讨论的表征理论的采纳。丘奇兰德夫妇倒是注意到仔细地避免返回到局域表征，并强烈地主张他们称作"矢量编码"的优点（P. S. Churchland and Sejnowksi 1992：163～174），但他们在没有充分接受分布式表征不再是该词传统意义上的表征的情况下停止了脚步。

这种情形也是与他们的分析中广泛运用的态空间描述混合在一起的。由态空间的特定的点来定义网络状态，导致了这样的印象：这样的点显得是唯独的。不幸的是，态空间描述并不能指明在种种参量之间的关系，也不能指明在子空间之间的任何关系。而且，我们必须想像高维空间时，我们的几何直觉也没用了。如同将隐含神经元解释成特征探测器一样，定义良好的动力学系统的态空间描述也可以是一种有用的工具。不过，当我们试图描述真正的复杂性时，态空间图有强化流行的原子论的倾向。

尽管他们坚持了表征，丘奇兰德夫妇还是在理论界成为认真到把联结论当作他们模型的中心成分的程度的第一批声音之一，从这个方面来看，他们是应该得到支持的。然而，无论是何种想像力都不可能把他们的趋法称作“后现代的”。

接下去，我希望在两个层次上考察分布式表征的含义。一方面，我希望表明，以这种方式来思考表征，对于神经网络在实践中的运用方式（作为分类器或是模式识别）是有借鉴意义的。这种讨论将更多地关注于技术问题，但是也为了表明——不只是陈述——后现代趋法的确具有实践后果，甚至涉及平凡的运用时也是如此。如果你已经相信这一点，并感到工程的例子也许不那么吸引人，那你可以越过下一节，直接进入关于分布式表征的一些哲学含义的讨论中[5]。

5.3 实际建模中分布式表征的重要性

在专门讨论神经网络之前，有必要交代一下传统的AI方法（如专家系统）[6]。由于它们是以逻辑关系的形式符号来运作的，因此采用的是传统的逻辑表征方法。这并不意味着它们本质上有什么错误，而只是意味着，你需要为所描述的领域建立非常好的形式模型。适

当的模型通常都要采用许多特设的规则来处理意外。特设的规则越多，模型就越加变得“分布式”。另一方面，最小神经网络就开始越来越像专家系统了。因此，在完全的定域分布式表征和完全的分布式表征之间，有着选择的范围。然而，在两者之间作出的选择，会强烈地影响对于问题的趋法，且如果我的这个分析是对的，同样会对结果产生影响。

5.3.1 分布度

当我们运用神经网络时，不仅所利用的网络的结构，而且网络中数据提呈的方式也会影响分布度。如果对数据进行了算法预处理以抽取或强调一定的特征，就会对数据强加上了特定的解释，网络的自由也就受到了约束。在大多数情况下，这可能是不可避免的，因为原材料往往杂乱而难以处理。但在进行大量数据压减和预处理以向神经网络提呈由预处理算法选取过的有了一些特征的数据的同时，你没有有效地运用神经计算能力。一旦获得了一些特征，任何的经典分类技术便可执行所需要的任务。进行深度预处理的进一步结果是，最小网络常常足以进行最后的分类。这些通常被训练到使种种神经元不是被激活到高水平就是低水平，即处于接通或是关闭，网络中的权重倾向于采取极端的正值或负值。结果是，网络最终由大量作出是非判断的逻辑门所构成，分布性被等效约简。

不幸的是，解决玩具问题的小网络被广泛运用于分析网络的行为。通常的原因是，只有小网络才允许被详细地分析。从我们的观点来看，这种论据是一种自满自足的预言：只有小网络被加以分析是因为只有小网络能够被加以分析；要避免网络具有分布式表征是因为它们不满足我们的框架[7]。运用模拟的数据并使用经过特意制作的网络以使这些数据成功地运行，这是一个相关的值得怀疑的

实际问题。这并不真正提高我们关于神经网络的能力和极限的知识——通过考察网络中分布式表征概念的富有意义的运用方式，有些事可以做得更好。

由于分布性不是一个绝对参量，而是取决于问题的语境和复杂性，也就不可能为它提供某种客观的定量标准。在缺乏定量的分布性定义时，一些定性的观察可以为我们提供帮助。首先，相互关联越是充分，网络中就越是倾向于分布。在此意义上，常规的后向传播网络——在各层之间有着充分的关联——就是够格的，假如其总体上有足够多的神经元的话。第二个考虑，涉及网络中的节点数。因为这同样是由所涉问题复杂性所决定的，因此就不可能给出一般性原理。不过，我希望表明，受过训练的网络是可以进行分析的，以决定是否大部分神经元已经移动到是非条件。在这种情况下，网络很可能太小了。因为在输入数据和权重值之间有着密切的关系，输入也就对此有着一定的作用。于是第三种考虑就必须涉及训练的样本数。如果只使用 4 个输入矢量来训练只有两种输出条件的网络[8]，那么肯定不会导致真正的分布。对一定问题适当的训练样本数，是由问题的复杂性所决定的。如果该问题非常简单，就没有必要使用神经网络。

5.3.2 多层感知机中的概括作用

神经网络成功地运行并非来自训练集的输入的能力，是其最重要的特性。网络能够在属于同类的训练样本中发现共同要素，并能在再度遇到这些要素时作出适当的反应。在设计一个网络时，优化这种能力是一个重要的考虑因素。

只要涉及后向传播网络，使用过它们的人都会确认如下的策略：所使用的网络不要超过解决该问题的需要，也不要过度地加

以训练。在这样的环境下，网络将建立起关于训练集的精确表征，使得在检测集中会将那些稍微不同于训练样本的样本忽略掉。这种考虑将使得实践中使用的网络尽可能小，但是，在良好的概括作用的条件与分布式表征的条件之间显然有一种张力。对此需要加以仔细分析。

神经网络的大小问题，开始受到了越来越多的关注[9]。希罗斯(Hirose)等人(1991)讨论了网络中隐含单元的数量。他们假定输入单元和输出单元的数量是固定的，然后发展一种可自动地变化隐含单元数量的算法。这可以缓解两个问题：第一个涉及胶着在局域最小值上，第二个涉及并不预先确知隐含单元的数量。

希罗斯等人的算法如下：在训练过程中，网络的均方误差 E 是受控的。如果经过了 100 次权重作用，E 的降低不超过 1%，那么就假定该网络的收敛或是遇到困难，或是它已经胶着于局域最小值，这就需要增加额外的隐含单元。一旦网络收敛了，过程就倒过来，即去掉隐含单元，直到网络不出现进一步收敛为止。于是，网络就被选定。（这种方法如同停泊汽车，把车开到直至听见了嘎吱声，然后退回来 6 英寸。）对此算法已进行了检验，一是采用 XOR 问题（参见注释 7），以及使得网络学习识别 36 个字，每个字由 8×8 个像素构成，每个像素仅仅取值 0 或 1。

我将它看作是对于神经计算的盲摸趋法(blind approach)，这是必须加以避免的。在我们目前的知识阶段，对网络遇上的困难应该给予大的关注，而不是以包揽了所有原理的“自动”程序来模糊对待，我在上一节中已经表达了这种反对意见。希罗斯等人在文章的结尾处(66 页)写道，概括作用的问题完全没有加以考虑，尽管他们还承认当隐含单元未以正确次序移除时他们陷入了困境(64 页)，但

是他们并没有考虑到这对于他们思考问题方式的含义。

概括作用肯定是西特思玛(Sietsma)和道(Dow)(1991)的文章所关心的问题。他们首先接受这样饱含智慧的说法：隐含单元越少越好。再加上另外关于一些神经计算的一般说法：网络的层数越多则概括作用越好；因此他们认为，长而窄的网络可以改善概括作用。这个假设用训练网络将一些正弦波分成3种频率的实验进行了测试。在此基础上他们加上了第4类，即代表了所有的不属于上述3类的其余频率。当输入层由64个单元组成时，比起通常用来解决XOR问题的网络——由5个神经元组成——这些网络的结果更接近于“分布的”。

随后的进程是对网络进行训练，并删除掉所有“冗余”单元以制造出最小网络。于是向其中插入额外的层，并对各种选择进行关于概括作用的测试。将噪声加入测试集以及训练样本中，并对其效果进行调查研究。他们有很重大的发现：

- **长而窄的网络的概括性不好。**这些网络进行的概括不仅不如其“母本”网络，而且当输入噪声时会产生较高的错误率(Sietsma and Dow 1991：75)。这种结果可以从我们关于分布式表征的讨论中得到理解。最小网络的串联导致的“逻辑”特性，将使得网络更加趋向于基于规则的框架，并因此而剧烈地降低了“分布”程度。网络中的表征越是“形象”(iconical)，概括作用就越是差。
- **向训练集中加入噪声会改进概括。**向输入矢量中加入“污点”肯定使得它们更加“分布化”，但是加入噪声的效应会带来另一个重要的结果：以有噪声的输入来训练的网络，不可能被删除到同样的程度。先前认为单元不会对于解有贡献，而现在证明是会有贡献的(Sietsma and Dow 1991：73)。这就强调了分布

式表征的核心特点，即特定节点的值没有意义，信息有赖于遍布整个网络中的权重模式。因此，增加隐含单元可以导致概括作用的改进(Sietsma and Dow 1991：68，78)。

此番对长而窄网络的研究结果同样可以推论到预处理上。在这种网络中，第一层实际上在施行某种特征提取，这类似于许多形式的预处理。这并不意味着，所有形式的预处理都必须避免：数据的化简常常是根本性的。然而，这将迫使我们仔细地考虑到预处理的效应。能够对良好的分布式表征有所贡献的信息不应该被排除掉，而且要时常记住，当从科学家所习惯的分析的、算法的框架角度来看待数据时，这样的信息也许不会被看作具有重要性。

5.3.3 改进概括作用

上面的观察能够导致对于神经网络的实际使用方式的某种改进吗？我认为，改变理论趋法，即将分布性看作某种优点而非缺点，可以对网络的性能产生影响。这里是从这种思路提出的一些实际考虑：

(i) **不要过度训练。**过度训练导致对于测试集的特定表征，而不是问题域的分布式表征。中断训练过程并对网络的性能用独立的测试组加以监控，可以防止过度训练。或者换一种方式，用确定网络是否收敛的标准来进行调节。

(ii) **向训练集中加入噪声。**加入噪声的效应已在上面讨论了，但应该记住，加入噪声并不会增加输入数据的信息量，它只会引起略微的动荡。在一些情形，如输入矢量的数量有限时，加入噪声可能是非常有帮助的，否则下一个考虑更为重要。

(iii) **加大输入样本的范围。**输入矢量应该尽可能有代表性。局

外物要尽可能包括进来。它们不仅仅具有如同加入噪声的效果，而且还向也许能使网络更好地分割特征空间提供了有意义的信息。

（iv）**增加输出种类的数量。**这种考虑，也许是最重要的考虑。正如索绪尔的语言分类系统，提供它们不是什么的样本可更好地定义范畴。网络可以拒斥的种类越多，在最后的分类中把握也就越大。加入一类，甚至更好是加入若干缺省种类，不仅将改进概括作用，而且也将使得系统在应对环境条件中的变化时更加鲁棒（背景噪声的水平，50 至 60 赫兹噪声的存在或不存在，动态范围，感知器的变化，减少错误，等等）。聪明地选取缺省种类（背景噪声，虚假种类，系统噪声，所有样本的共同因素，等等），将允许我们不扰动内部就对网络进行调整。例如，为所有样本的共同因素（如背景噪声）提供独立的类别，不仅将向网络提供额外的选择，而且由于它必须建立关于此共同因素的独立表征，并将其与真实的种类区别开来，这就使得真实的种类的表征更加鲁棒。对输出种类的创造性选择，也容许我们以同样的数据而做不同的事情。

（v）**运用真实数据。**除了运用模拟数据常常导致自回答（self-answering）问题之外，运用真实数据会使得分布式表征的若干重要因素自动地结合起来。噪声和变化自动进来了，也不容易受到在问题和解答之间插入某个“大理论”的诱惑。

（vi）**仔细地考虑预处理和后处理。**不要从数据中预处理出特性。宁可对网络输出的趋势和错误进行分析。

后处理是还没有得到足够重视的神经网络的一个方面，因此我将

在本节中对此问题进行若干评注。如果网络的输入和网络自身被制备成尽可能地“分布”，那么这也会影响输出的本性。首先，现在可以有多于特定解所需的输出神经元。其次，较少的预处理，特别是对于输入矢量的较少的平均化，意味着必须处理的是具有更广分布的大量输入。尽管这可能会增加输出的即时错误率，但辅以适当的后处理，网络的运行平均说来将更好。

适当的后处理可能包括对于输出的平均或过滤，寻求“好的”分类的更好标准，或使用输出作为随后的处理阶段的输入。通用的格言看来是，不要在网络的输入处做那些可以在输出处可做的事。

5.4 表征的解构

在本节，我将转到关于表征讨论的哲学含义，特别是转到分布式思维与后结构理论方面的关系[10]。后结构主义是一种重要的、激动人心的理论趋法，向许多关于语言和科学的传统趋法的核心假定提出了挑战。我希望指出，对复杂系统建模的分布式趋法，与后结构主义有许多密切关系，后结构主义能够为我们的科学和技术实践提供至关重要的问题。反过来，在运用这些后结构趋法的实践中所获得的成果，检验了反对把后结构主义仅仅看作是一种具有解构倾向的哲学分支的立场。

后结构思想的重要特点之一，特别是在德里达的著作中，是否认语言的透明性。语词的意义是词代表其本身时的“表现”，这种思想导致了决定性意义的幻象。一个具有特定意义的语词的鉴明，是德里达称作“表现的形而上学”的实例。他认为，不存在语词与其意义之间的一一对应关系。意义是系统中所有的词(或不如说是所有的语符)的交互作用的结果。它是系统中动力学的效果，而不是系统

组分与世界中对象之间的直接关系。这并未否认世界与系统之间的所有关系。恰恰相反，系统的成功主要取决于系统与其环境之间的相互作用的有效性。这里的确否认的是，当我们处理如语言或大脑这样的复杂系统时，这些关系能够以决定性术语来揭示。复杂性如果不丧失正好是我们感兴趣的系统之能力，就不可能被简化成直接关系。这样的能力是作为系统中组分之间的非线性、分布式关系的结果而出现的。

尽管我们不可能以决定论术语来描绘复杂系统的本质，但是我们也不可能阻止甚至无法避免建构起某种对于一定时刻系统本性的解释。然而，这些解释在原则上是有局限的。我们总是被限制在关于系统的简单印象中。这些印象总是从某个角度获得的，揭示了某一时刻系统的某些方面。没有什么可以阻挡我们尝试解释系统——我们可以获取尽可能多的图像——只要我们意识到每一特定图像的局限。由于复杂系统处于不断变化之中(即不处于平衡之中)，也就不可能如同在拼图游戏中把片段正好置于其正确位置上那样将一系列图像连接起来。我们可以通过并列、比较、拼贴，把它们结合成序列以发展成某个叙事情节，并由此以也许更有创造性的方式建立我们对于系统的理解。其间有这样的危险，即陷落在拼贴某种特定图像中，并认为这是它的一个优选位置。由于它不仅仅否认特定角度的局限性，而且也阻止了进一步的探索，所以这种拼贴必须被打破，被不停息地显示从一个角度的固定边界所导致的矛盾打破。指出伴随着此种封闭而来的矛盾，是一种被德里达称作“解构”的活动。

目前为止的讨论，实际上是在没有将其塞入后结构话语的情况下对表征概念的解构。采用这种趋法的动力之一是希望表明，后结构主义不只是一种话语分析的形式，其也构成了一种对于科学和技术具

有实践意义的理论立场。另一个动力，是需要尽可能地抵制解构话语的诱惑——避免过于轻率地滑入其中——因为这里总是存在着风险的，且这种风险常常为哲学家们所低估，以致在不知道对手是谁的情况下进行比赛。然而，由于表征概念具有重要的哲学意蕴，我们必须搞清楚分布式表征这个分支。

5.4.1 分布式记号语言学

任何表征理论，都是关于语符本性的特定理论的结果。在大多数记号语言学系统中，语符是通过指称一定事物而获得意义的，它表征着指称物。索绪尔(1974)通过主张语符的意义是其对系统中所有其他语符的关系的结果，而提出了一种分布式记号语言学。因此意义不是任何离散单元的特殊特征，而是有差异的系统的结果。为了形成语符的意义，不仅要考察那个语符，而且要考虑它所牵涉的整个系统——意义因此是分布式的。然而，由于他坚持在能指和所指之间进行区分，他的系统仍然是表征的。尽管他以分布式的方式来运作系统，他并不希望放弃塔桥图中的上层。正是在这一点上，德里达展开了对于索绪尔记号语言学的攻击。

在一个分布式记号语言学系统中，语符是由它对于其他语符的关系之和所构成。德里达将这种任何两个语符之间的关系称作“痕迹”。虽然是这样，痕迹本身却不具有可以明晰的构思性内容。它在语符本身的层次上运行，而不是在高于或低于语符的元层次上运行。以这种方式理解，痕迹就等价于神经元系统中的权重。网络中节点的意义，不是节点本身的某种特征的结果，它是经过权重后的输入和输出的模式的结果，这一模式将其中一个节点与其他的节点相关联。权重，正如同痕迹，并不代表任何特定的东西。

为了提供记号语言学的一般理论以语法学“科学”，德里达明确

否认表征理论。他解构了此塔，结果仅仅给我们留下了材质的底层。系统不再受形而上学的外在所支配，也不受发布命令的内部中心所支配。它只是由网络中痕迹的分布式相互作用构成(Derrida 1982：3～27)。而且，这里的引文中出现的德里达的延异概念，也是指在联结论模型特别是归复型神经网络中展现其本身。这种类型的网络具有许多的反馈环：某节点的输出可以成为该节点本身的输入，其间可以是通过其他节点，也可以是不通过其他节点。节点的活动因此不仅是由它与其他节点的差异所决定的，也是延迟到其本身的(以及其他的)活动反射回自身后决定的。在这种相互作用的复杂模式中，不可能说某种语符(或节点)表征了某种特定事物。一个强的表征理论，将总是强加上了表现的形而上学。它实际上主张了两种系统——语符本身以及外在于它们的语符的意义，通过表征过程而相互呈现。一个分布式的记号语言学理论使得这种划界有问题。它再次主张，没有东西可以脱离能决定痕迹的语符系统的范围，因为"外部"本身不可能摆脱痕迹的逻辑。是否应该从外在找到痕迹的来源呢？这将面临着同样的分裂、运动和抹除。在"广义参考物"中，外在和内在是相互为依据的；我们这里得到了限于能指层次的分布式表征。德里达(1982：24)的阐述如下：

> 由于痕迹不是表现，而是表现的模拟物，表现脱离了自身，置换了自身，依据自身，确切地说它没有存在之处——抹除了属于它的结构……在形而上学的语言中的这种结构的悖论，是形而上学概念的倒置，这产生了如下的效应：呈现成为语符的语符，痕迹的痕迹。归根结底，它不再是任何参考物所指称的东西。它变成了广义参考物结构中的功能。它就是痕迹，痕迹抹除的痕迹。

我们倾向于某种一般性表征理论是容易的，也许可以由我们文化中形象的重要性来解释。“我们都是优先‘通过眼睛’来感知的，这不仅是为了行动，而且甚至在眼前没有实物的情形时亦如此”(Derrida 1983：4)。当我们说，一幅形象胜过千言万语——意味着一幅形象无论如何要比语言更有说服力——我们成了表现的形而上学的俘虏。我们相信，一幅形象以某种外观承载着意义，这就脱离了由分布式记号语言学所描述的举荐蓝本(refrral)的作用。文本也许必须得到解释，但是形象却能自己诉说或我们都相信是如此。这个概念，受到后结构理论的强烈抵制。反对表征的一种证据，同时就是对于形象本身的交织本性的证据(Derrida 1976：36)。

当我们否认表征理论的可能性时，涉及分布式系统与世界之间关系的问题却并没有消失。从联结论角度，我们可以尝试给出一种回答。表征系统中，表征和被表征的东西运行在不同的逻辑层面上；它们属于不同的范畴。而神经网络中，情况不是这样。在这里，进入到网络的感觉痕迹与网络中相互作用中的痕迹之间，种类上是没有区别的。在一定的意义上，我们获得的是外部世界在内部世界中的重复(repeated)或归复(reiterated)，由此解构了在外部和内部之间的区分。于是两者之间的鸿沟便消失了[11]。

表征理论，结合我们对形象的偏爱，最终构成了一种自我陶醉。这种自我陶醉，是由我们试图通过使用抽象理性和加强表征理论在计算机上模仿人类的智慧所成就的。

5.4.2 机器概念

我们常常把机器设想成似乎含有某种非人化的东西。它们至多被看作是有用之物，甚至本质上就是工具；最糟糕的情况，则是被看作灵魂的破坏者。正是这种双重关系——依赖和憎恶——导致了试

图在无生命的机器上复制生命智能的一种奇怪的态度。在许多人看来，这是现代科学的最终追求，这是最后的大餐。而另一些人认为，它成了把人当成工具理解的基本信号，这种趋法的假设是对人的价值的重要性的否定。一方面，这两种感情并非是截然分开的：第一种主张，是要使得人之所以成为人的本质东西可以在机器上呈现出来；而第二种主张，害怕的是一旦这成为现实那么可以称作人的标志的东西也就没有了。两者都被还原成机器的思想所迷惑了。

我们希望在机器上模拟的称作"人"的东西究竟是什么？我们希望模拟的本来究竟是什么？如果我们希望表征一个人，我们必须相信，我们已经有了一种关于什么是人的理论。一些这种理论的色彩(以表征为外观)在选择模拟物的基本属性中展现出来——一种抽象的、雌雄同体的智能被明确叫做"人工的"。这种项目所要表征的是模拟物内部的东西，而其真正形式却只可能从外部发现。但是，痕迹的"逻辑"却打乱了表征(内部)和被表征(外部)两方面。当内部的封闭被打破时，我们发现了一种不同的拟态，它由相互定义的反身过程构成(参见 Clark 1985)。内部与外部变得交织起来。一旦我们放弃对于真正的、可被忠实复制的真实事物的怀旧情绪，我们就进入了鲍德里亚(Baudrillard 1984)称作的"超现实"(hyperreal)中。当我们说内部与外部是相互交叠的，这不仅意味着我们必须理解复杂系统的动力学，而且还意味着前面所主张的对于实在的理解。因为"实在"(reality)主要是以语言的方式被触及的，我们对于世界的理解也就是遵从于痕迹的逻辑。世界不再是简单的意义之源。技术(计算机)和媒体(特别是电视)在我们关于世界的理解方式中的媒介作用不断增强，使得这种事态加剧了。我们面临的，不再是现实，而是现实的模拟物，即一种超现实。鲍德里亚和埃科(Umberto Eco)(1987)两

人——特意参照了美国的后现代文化——已对这些概念进行了分析。

鲍德里亚(1984: 254)解释道，向超现实的移动是“对所有参照物的清算”的结果。我们的语符系统，不再锁定在现实世界；它们漂浮在我们的银幕上，繁殖于我们的计算机和数据库中。也没有了“真实的”和“想像的”(对现实的意义解释)；所有的东西都瓦解在语符及其相互作用的层面上——“模拟物的运动”。他是这样叙述的：

> 不再有仿造，也没有了复制，甚至也没有了悖论。更确切地说，问题是以现实的语符取代了现实本身——也就是，一种由其运行两重性而阻止了所有现实过程的操作，一种提供了全部的现实的语符并缩短了其所有的兴衰变迁的亚稳定的、程序化的、完美描述的机器……超现实因此遮蔽了想像力，只留下轨道式模型重现和模拟差异生成的空间。(254 页)

作为对表征概念的替代，鲍德里亚选择了“模拟”(simulation)概念。模拟并不试图表征对某种现实的本质性抽象；而是试图重复它，以此削弱现实和被模拟物之间的区别。

> 不同于表征试图通过将其解释为虚假表征而吸取模拟，模拟将表征的整个大厦包蔽起来看作模拟物。这些可看作是图像的各个连续阶段：
>
> - 它是基本现实的反映
> - 它标志着并曲解着基本的现实
> - 它标志着基本现实的缺席

● 它与任何现实都无关：它本身就是纯粹模拟物。（256 页）

让我以实际例子来说明。神经网络通常是在数字计算机上“模拟的”。计算机本身并不包含相互关联节点的网络，但可设法使其表现得如此。现在，将麦克风与计算机关联起来，藉此可以向“网络”提供信息，并运用计算机的输出来产生某些效应，如接通某个电机。这种系统可以执行某种任务（例如，当汽车接近时打开一扇门）——与神经网络所具有的特定能力（将此汽车声与其他噪声区分开来）类似的能力。现在有可能提出我们是否在与模拟作用共事的问题。然而，要点在于，答案无关紧要。因为系统的行为如同网络在进行此项工作，于是在“真实的”网络与“模拟的”网络之间的区分消失了。

这些思想，对于这里考虑的问题——人和机器之间的区别以及对于机器智能的疑问——具有重要的后果。这种疑问也许可以作为表征理论的结果而消亡。这种理论作了形而上学的假设：人的某种本质的东西必须以抽象的方式来表征，而不是使人认识到，遵从模拟物的逻辑，我们得到的只是必须反身地反复或重复。问题就从“如何在机器上模拟人？”转变成了“何种机器才是往往被称之为人的那种存在？”

正是这个问题，从早期的《方案》（Project）（1895，身后于 1950 年出版）到著名的“关于‘神秘书写板’的注”（Notes on the “Mystic Writing-Pad”）（1925），一直困扰着弗洛伊德。他试图从痕迹和机器的比喻来设想心灵（psyche）（Bass 1984：77），而不求助于抽象认知概念的运用。《方案》中的网络模型，以及简单的被称作“神秘书写板”的机器，最终把某种对于“机器”的理解与某种对于“书写”的

理解结合了起来。不过，对于弗洛伊德来说，“书写板”仍然只是一种比喻；它“表征”心灵的工作。德里达(1978：199)在对弗洛伊德的研究中敦促我们不仅仅是运用比喻，而且要超越它：

> 心灵的内容将由其本质上是不可还原的图标的文本来表征。心灵装置的结构，将由书写机器来表征。这些表征强加于我们的问题是什么？我们不必追问，书写机器——例如“关于‘神秘书写板’的注”中描述的书写机器——对于表征心灵的工作是否是一个好的比喻，还是仅为了表征心灵的书写而必须创作的装置；以及我们必须追问，在机器中设计的和释放的同心灵书写的东西的仿造物意味着什么。以及若不然，心灵的确是某种文本则不然，但是：文本是什么？如果心灵能够被文本所表征，那么心灵必定是什么？如果机器和文本都没有心灵的起源，那么就不会存在没有文本的心灵的领域。最后，在心灵、书写和空域之间必须由什么样的关系才使得这种比喻转移成为可能，这不仅仅也不主要是在理论话语之中，而是在心灵、文本和技术的历史之中。

德里达在此指出，弗洛伊德停止于此是缺乏“对于表征本身、或心灵和机器之间‘相似’的含义的任何考虑”(Clark 1985：310)。这使得他无法意识到心灵不多不少正好是“书写机”。机器概念对于德里达是重要的(Bass 1984：77)。在去除了关于语言和大脑的形而上学维度之后，剩下的是比较纯洁的物质机器[12]。机器不仅在具有复合的结构和部分的意义上是复杂的，而且在以真正分布的方式经历了无尽的循环、转换和交互作用的意义上也是复杂的。

它承载着循环并以此作为反驳表征的论据，但这完全不是反科

学的(anti-scientific)。它仅仅是一种反对特定科学策略的论据，此种策略假定复杂性能够被还原成特定的特征，然后在机器中得以表征。相反，它是一种关于复杂性本性的适当论据，某种也许能够在机器中“重复”的东西，机器自身也必须充分复杂以处理复杂性的分布本性。

我们的技术能否制造出这样的机器，仍然是悬而未决的问题。不过，业已清楚的是，计算机在这些发展过程中仍将扮演某种悖论的角色。一方面，它是试图执行表征的形而上学的理想工具。另一方面，它又可以成为分布式书写机，“如同车轮”靠自身运行。

让我们重复一下以上的论证，因为它对本书至关重要。基于形式符号系统的模型，使得经典的表征理论成为其基本的内容。表征的主要问题，在于符号与其意义之间的关系。建立这种关系的方式有两种。一种可以认为，此种关系是“自然的”、以某种先验的方式被决定；另一种可以认为，它是由决定着这种关系的外部设计者所建立起来的。第一种选择是一种强烈的形而上学的选择，因为它主张意义是由某种基本的、包揽一切的定律所决定的。此种趋法在这里必须要加以排除，因为我的论证的主要支柱是，对于复杂性的理解不需要借助形而上学的基础。第二种选择——在此关系是由设计者决策的结果——是可接受的，只要可能假定存在着某种活跃的、外部的行为者。当充分掌握了信息的建模者对于勾勒好框架的系统进行建模时，情形即是如此。不过，当我们处理具有高度复杂性的自主的、自组织的系统时，这第二种选择也成了形而上学的选择。只要我们放弃表征概念，这些形而上学问题也就消失了。

我们仍然可以有能够从其环境编码信息并以此为基础运行的机器，而不必将其内部机制聚集成表征模糊块。这并不意味着，这种

非表征的、分布式的系统，其内部结构将是某种继续的、同质的整体。它也可以分化成若干种功能的子结构，但它们仍然是功能单元，而非表征单元。大多数这样的单元，将很可能具有分布式结构，并将有与其他单元的高度关联。当我们在下一章转向自组织时，内部结构的发展将被详细分析。在这里，我们将只是补充：联结论模型是否终将完全地模拟语言和大脑也是没有保证的。不过，随着对形而上学基础的拒斥，此问题成为一个偶发性问题，不可能预先给予答案。

5.4.3 表征对科学和理论关系的影响

在转到自组织之前，我将扼要地讨论分布式理论对于科学理论状态的一些影响。科学的中心价值之一是具有客观性。只有假定科学家可以将事物看作如同就是那样之时，客观理性原理才是起作用的。然而，追求客观性不可避免会引入科学家(作为认识主体的一方)与研究的客观性(在某种纯粹形式中呈现的另一方)之间的裂缝。在一篇论大学的作用和地位的论文中，德里达(1983)较详细地分析了这个问题：

> 现代占支配地位的理性原理，必须与对作为对象的事物的本质作出解释联手而行，后者是一种作为表征(*Vorstellung*)而存在、摆在并安置于主体之前的客体。这主体，即一个说“我”的人，是肯定自身的自我，从而保证了他对于全部所是的技术把握。*repraesentatio*（表征）的“re-”也表达了一种运动，该运动说明了——“提供理由给”——在场之物是通过描绘其表现和将其带到了表征主体的面前和自我理解而遇到的。……但确是事实的是，在海德格尔的意义上，一个表现力强的人的讽刺画，容易赋予他一双

无情的双眼——永久地盯住他要主宰的自然，必要时就通过将它固定在自己的面前，或是如同捕食之鸟猛扑的方式掠夺。(9～10 页)

放弃表征的理论因此就意味着对于科学与理论之间僵硬边界的解构。如果真正的客观性不再可能，那么理论便贯穿于科学活动所有层次之中。理论与实践之间的明确区分的丧失，不仅将导致科学衍生物，也将出现理论衍生物。理论不能再为自身而进行，不能再为了纯粹知识而追求。

德里达认为，这在过去是可能的，在过去的某些地方曾相信“纯粹”科学，未受到来自国家的压力和经济的约束。而现在这是不再可能了，或更准确地说，实际上从来没有过。

在每一个这样的领域(理论物理学、天体物理学、化学、分子生物学，如此等等)——它们比以前更加交织起来了——所谓的基本哲学问题不再简单地采取抽象的形式，认识论的问题往往是产生于事实之后：它们以最广泛的方式产生在科学研究的核心部位。人们不可能再将技术的一方，与理论、科学和理性的另一方之间进行区分了。(12 页)

与此同时，科学丧失了透明的客观性，哲学丧失了避免偶发性的奢望。这将是后现代性的持久特征，即科学家和哲学家都丧失了他们的纯洁性。因为一定的表征理论就蕴涵着一定的意义理论——而意义是我们赖以生活的东西——我们对这样一种理论的选择就具有了重要的伦理含义。我将在最后一章回头来谈这个问题。

注 释：

1. 生产力与结构也是福多尔和佩利夏因(Fodor and Pylyshyn 1988)提出来反对联结论的论据的两个口号。 这在第 2 章中进行了讨论。

2. 特定神经元的转移函数也可以是不同的。 这在发现经济的执行中具有实际重要性，但是在理论中，如要实现转移函数的变化，可以通过加入偏好术语，或者如有必要就采用若干额外的具有类似于其余的转移函数的神经元。

3. 当然，如果网络具有自组织能力，丢失的神经元的功能可以被其他的所代替，而不必从缺痕中建立起新的表征。 自组织将在下一章讨论。

4. Sterelny (1990: 187)列出了 6 个"严肃的问题"。 其中一些相当类似而可以组合起来。 我也将讨论它们，但是以不同于他所提出的方式。

5. 严格说来，我们不应该在这种语境中再使用"表征"这个词。 在不引入新的和可能引起混淆的术语的情况下，我将继续使用这个术语，请记住，"分布式表征"这个术语包含着内部矛盾。 在海德格尔传统中，人们也许已经使用了"分布式~~表征~~"(distributed ~~representation~~) * 。

6. 专家系统是知识域的形式模型，运用规则集在计算机上运行。

7. 一个经典的例子是所谓的异或(XOR)问题。 它被用作关于分类器的传统测试案例，因为其构成了一个无法线性分离的问题。 由 6 个神经元构成的神经网络可以解决这个问题。 我并不认为对这种网络的分析能对我们理解神经网络计算有许多贡献。

8. 这些参数适用于 XOR 问题。 参见上面的注释 7。

9. 例如，在 ICANN - 92———一个大型的神经网络会议(参见 Aleksander and Taylor 1992)——有几篇文章讨论了这个问题。 Judd(1992)表明，宽而浅的网络是计算上有效率的；Hecht-Nielsen (1992)指出，对于解决复杂问题所必须要的高维网络，收敛的困难要比人们想像得小。

10. 这里提供的哲学论据，肯定不是反对表征的唯一的论据。 例如，关于表征的问题，是罗蒂(Rorty 1980)的《哲学与自然之镜》(Philosophy and the Mirror of Nature)的中心论题之一，他在书中也与乔姆斯基、福多尔(244～256 页)论战。 Joseph Rouse (1996: 205～236)从戴维森视角来反对表征。 Norris (1992: 59～82)的分析试图表明在戴维森与德里达之间的相似性，并作出结论说，这种比较不是一个简单的比较。

11. 这个问题被 Schalkwyk(1991)详细地讨论过，讨论中结合了来自维特根施泰因和德里达两方面的见解。 语符的可重复的结构必然性，是塞尔与德里达(1988)之争中最重要的中心问题之一。

12. 人们也许会争辩道，鲍德里亚和德里达两者都希望摧毁表征的塔桥图，但却是以不同的方式。 鲍德里亚要更为"后现代"，是让此图像崩溃到其顶层——所有的事物都是模拟的，而德里达希望让其崩溃到其底层——所有的事物都是材质的。 如果人们落入到只有一层的思想中，也许差异也就没有那么大。 不过，人们能够看到，鲍德里亚的趋法是如何导致了"怎样都行"的后现代主义风格；而这对德里达关于文本状况的理解来说是某种陌生的东西。

* 原文 146 页如此。 ——译者

第 6 章　复杂系统中的自组织

到此为止，我们已经对复杂系统的结构特征给予了大量关注。在本章中，将关注结构是怎样产生、发展和变化的。“结构”概念适用于系统发展起来的内部机制，这样的机制一方面用来接受、编码、转换和贮存信息，另一方面通过对某种形式的输出而对这样的信息作出反应。这里的主要任务是要表明，在没有外部设计者的干预或不存在某种中央化形式的内部控制的情形下，内部结构也能够发生进化。如果系统的能力满足一些约束条件，则能够通过自组织过程而发展起分布形式的内部结构。这种过程中，结构既不是对于外部的被动的反应，也不是主动的、预先编程的内部因素的结果，而是在环境、系统的现存状态和系统的历史之间的复杂相互作用的结果。

贯穿整个西方智力传统中的大多数哲学立场，几乎都对秩序和结构的自发涌现（*spontaneous* emergence）持怀疑态度。由于对此种涌现缺乏合理解释，通常就假定了某种组织行为者——上帝（作为终极设计者）或某种其他的先验原理。而自组织既非某种神秘过程也非某

种随机过程，不应该与我们任何的正常感受对立起来。这也是我希望在本章中所表明的。

随后将讨论若干自组织系统的例子，这里一个简单（也非常有限）的例子可有助于引入自组织的基本思想。考虑水库中的一群鱼，假定我们可以通过鱼群的大小来测量其生存状态。鱼群的状况将取决于大量的因素，包括食物的丰富性、水的温度、氧和光线的充裕情况、鱼龄的大小，等等。随着这些条件的变化，鱼群将调节自身的大小以适应基本的生存条件，尽管每一条鱼只可能照顾自己的利益。这个鱼群系统作为一个整体，将自身组织起来以保证在系统与环境之间的最好配合。在鱼群将根据过去的经验对变化着的条件敏感的意义上，这种组织过程也是一个适应性过程。在此没有行为者去决定应该发生什么，而且也不是每一条鱼理解了情形的复杂性。鱼群的组织活动，是作为系统中种种要素与其环境之间相互作用的结果而形成的。

一个简单的引介性评注可能是必要的。下面要说明的自组织基本特征和原理是相当普遍的，当然这与我们前面第1章关于复杂系统的一般特征有些重叠。它们都是从大量不同的事例中提炼出来的，因此，所有这些特征不可能呈现于每一事例。不过，系统越是复杂，这些特征就越是明显。我们的目的是要理解自组织动力学复杂系统的普遍特征。

6.1 若干关键方面

6.1.1 探讨自组织的一种定义

面对定义复杂现象的困难，我们将为自组织提供一个工作定义，并通过一些特征和例子来加以阐明：

> 自组织是复杂系统的一种能力，它使得系统可以自发地、适应性地发展或改变其内部结构，以更好地应付或处理它们的环境。

大脑是我们感兴趣的这种系统的极好的例子。在一定的约束条件下——包括物理的、生物的和遗传的——大脑不得不发展起对于环境的理解以在环境中有效地运行。可以合理地认为，大脑从其起源(*ab initio*)便包含着可以应付可能发生情况的所有程序，因此我们可以有把握地假定，大脑已经具备进行学习的能力。使得大脑能够产生记忆已学之事所必须的结构变化，因此必定是自发地形成的。

同样具有自组织性质的不同系统，未必显示出同样的特征范围。一个活细胞肯定可以归类于自组织系统，但是其内部结构与一个国家的经济系统相比要稳定一些。一个经济系统，在其变化着内部结构以对大量因素(货币供应、增长率、政治稳定性、自然灾害，如此等等)作出响应的意义上，是自组织的。尽管所有这些相互作用过分复杂，难以建构起决定论模型，但是对系统结构的大规模的干预也是可能的(货币升值，利率调整，如此等等)。不过，这些干预的效果只是短期可预见的，因为系统的自发调整涉及太多因素的复杂相互作用——其中许多因素是完全无法控制的。

自组织系统的另一个例子是语言。为了进行交流，语言必定要具有可辨识的结构。为了能够在极其不同的环境中保持其功能，结构必须是可以调节的——尤其是涉及意义时。因为许多个体在运用同一种语言，所以这些调整不可能仅仅发生在个体决策的层面上。变化来自大量个体的相互作用[1]。社会相互作用系统，即文化系统，与语言系统有许多共同的特征。这里建立的复杂系统的模型，对于语言理论，以及对于一些更一般的哲学和伦理学问题，都具有一定的

意义。这些将在最后一章中加以讨论。

这些例子表明，自组织可以在不同的层次上、根据不同的约束而运行。尽管在不同的复杂系统中例子是有差异的，然而自组织过程具有许多一般特征，现在我们就要对此进行讨论。

6.2 自组织系统的属性

尽管在多种自组织的、具有不同功能的复杂系统之间有重要的差异，但是复杂系统的一般模型框架却具有共同的属性。正如我们在第1章中指出的，一个复杂系统由大量简单单元构成，这些单元形成了以高度非线性关联起来的网络中的节点。系统的行为主要不是由系统的个体组分所决定的，而是复杂的相互作用模式的结果。自组织系统的一般属性包括如下一些：

（i）系统的结构不是先验设计的结果，也不是直接由外部条件所决定的。它是系统与其环境之间相互作用的结果。

（ii）系统的内部结构可以动态地适应环境的变化，即便这些变化没有规律。

（iii）自组织不只是如同反馈或可以线性描述的调整的结果。它包括了较高级的秩序、非线性的过程，不可能以一组线性微分方程来进行建模。一个以接通或关闭的方式对其环境进行反应的自动调温计，不是一个自组织的例子。

（iv）自组织是系统作为一个整体的（或充分大量子系统的）涌现性质。系统的个体组分仅仅依靠局域信息和一般原理而运行。宏观行为从微观的相互作用中涌现出来，微观相互作用本身只有非常微弱的信息量（仅仅是痕迹）。将分析限制在微观水

平上，有可能按照一些简单转化来解释每一要素的行为。从宏观进行观察，简单的、局域的相互作用可以导致复杂行为。

（v）自组织系统的复杂性能够增长。由于它们必须从经验中“学习”，它们必须“记忆”先前遭遇过的情形并将之与新的情形进行比较。如果更多“先前的信息”可以被存贮，系统将能够进行更好的比较。这种复杂性的增长，意味着熵的局域倒转，这对于能量或信息在系统流通是必要的。复杂性的增长，对于为何自组织系统倾向于变老，也可以给予部分的解释。由于这些系统受到物理世界的一定约束，它们不可避免地在某个点达到饱和。

（vi）没有某种形式的记忆就不可能有自组织，这与上一点密切有关。没有记忆，系统就不可能比仅仅作为镜子对环境进行反映做得更好。因此，一个自组织系统总是具有**历史**的。这种历时性因素在对于系统的任何描述中都是不能忽视的，因为系统先前的状况会对现在的行为产生至关重要的影响。另一方面，没有某种选择性遗忘，记忆也是不可能的。仅仅是堆积信息而没有某种形式的整合，会使它没有意义。整合不是通过系统中某种形式的决策而“实施的”。没有使用的信息就会消退。这个过程不仅为记忆创造了空间，而且更重要的是，还为所贮存的模式的意义提供了度量。一件东西被运用的次数越多，在记忆中的“表征”就越是强烈。要么利用它，要么丢掉它。只有系统能够记忆**和**遗忘，自组织才是可能的。

（vii）因为自组织过程不是被特定目标所指引或决定，对这种系统的**功能**进行讨论往往是困难的。只要我们引入了功能的概念，我们就有了拟人性或是为系统的结构引入了某种外部理

由的风险，而这些正是我们力图避免的。但当一个系统在更大系统的语境中被描述时，对子系统功能的讨论只有在那种语境中才是可能的。当我们说一头狮子的内分泌系统的“功能”时，是以这头狮子为参照的，但是这样就难以同时谈论狮子本身的功能了。我们可以谈论生态系统中捕食者的“功能”，但不可能谈论此生态系统的功能。功能概念与我们对于复杂系统的描述密切关联。自组织的过程不可能是由实施某种功能的企图而推动的；更确切地说它是某种进化过程的结果，系统如果不能够适应更复杂的环境，便不可能生存。

(viii) 类似地，对自组织系统以粗糙的还原主义加以描述是不可能。因为微观单元并不“知道”大规模的效应，同时这些效应是以集体的方式展示出来，不包括任何除了这些微观单元之外的东西，因此对系统的种种“层次”不可能独立地加以描述。各层次原则上是相互交织的。然而，对于描述涌现性质时运用还原主义的抵制，并不意味着对于唯物主义原则的任何放弃。

简而言之，复杂系统中的自组织过程以如下的方式运行。从外部世界来的信息簇流入到系统中。这种信息将影响系统中的某些组分的相互作用——它将改变网络中的权重值。按照赫布规则(第1章中进行了讨论)，如果一定的簇是规则地呈现的，系统将获得一组稳定的、“表征”着那簇(即在特定群簇每次呈现时导致特定的活动模式)的权重。如果两个簇是规则地共同出现的，系统将自动地建立两者之间的某种联系。例如，如果一定的事态规则地引起对于系统的损害，系统便将联系那种有损害的条件，而不必预先知道那种条件是有害的。随着系统遇上环境中的不同条件，它在由系统可利用的记

忆量所决定的约束下，产生出新的结构以“表征”这些条件。这种过程可以从数学上加以刻画(Grossberg 1987，1988；Kauffman 1993；Kohonen 1988)，但是它与弗洛伊德(Freud 1950)的关于大脑如何发展起其结构的神经学模型没有什么原则性不同。

现在还需要讨论的是系统中的相互作用藉此进行调整的实际原则。

6.2.1 自组织的基本原理

自组织可以用一种以上的方式进行建模，但是大多数模型都是采取网络中相互关联的简单处理单元的系统。我将集中关注神经网络模型[2]。这里扼要复述一下，神经网络由一些简单的通过突触相互关联起来的神经元构成。这些突触具有不同的强度，意味着神经元以某种复杂的、非线性的方式发生相互作用。此种系统可以非常形象地展示成相互关联的节点的网络，其中每一相互关联都有一定的强度或“权重”。所有的节点基本上都是类似的，网络的行为由权重值决定，而权重值可以被调整。决定着调整的，是以所涉及节点的局域可利用信息为基础的简单规则。一条规则(赫布规则)是，如果两个关联着的神经元被激活，那么相应的权重值便会增加。以这种方式，网络可以发展起以关联的动力学结构为基础的活动模式。

但是，系统的结构是如何在对周围环境的条件作出响应中发展起来的？只有信息可以从外部进入到系统中，这才是可能的。因此，至少一些关联必须在感受器或传感器处停止，从而感受环境的某些方面，并相应地激活系统[3]。在环境中的某种事件现在可以引起系统中的某种活动，这种活动可以用来转变此系统的结构，而且只是藉每一节点可利用的信息——并不需要某种全局性视野。在信息并没有以对称方式送进同质网络的条件下，网络的节点将被不规则地激活。

一些神经元组将会比另一些更加活跃。只要通过增加伴随着活化节点的权重，并减少其余的权重，这种活动模式就会被加强。如果外部的事件并不再次出现，这种模式最终将逐渐消失（被遗忘）或被其他模式所代替。如果事件是有意义的，在它经常出现的意义上，此模式将在此事件每次出现时得到加强。以这种方式，系统通过自组织过程发展起稳定的结构，使它能够识别重要事件。

因为自组织的最重要方面是结构的涌现，是通过并没有进入到全局模式的微观单元活动进行的，所以决定着局域的权重和节点行为的原理是非常重要的。下面的清单中提供了任何系统中自组织的一些先决条件，它们对于相应的理解是基本的[4]：

（i）系统由**大量的**微观要素或单元组成，它们起初是相对未分化的，即在此不需要预定的结构[5]。在神经网络的术语中，这意味着网络始于所有权重具有的随机值。

（ii）关联强度只是作为**局域信息**的结果而变化。这些变化常常是自维持的（包括了正反馈），并引起系统从未分化状态开始发展。

（iii）单元之中存在着**竞争**。争夺有限的资源，是结构发展背后的基本驱动力。较强单元以其他单元为代价而兴旺。如果资源无限（即如果可以无限制地生长），那么便不会导致有意义的结构进化。边界、限度和约束是结构的先决条件。

（iv）至少在一些单元之中存在着**合作**。如果只有某个单元是获胜者，那么所形成的结构对于自组织的进化便会过于简单。对于模式之间形成联合，合作也是必要的。相互加强和合作，对于丰富的、有意义的结构是先决条件。

（v）单元之中的相互作用必须是非线性的。小的变化必须能够引发大的后果，模式的组合应该可导致新模式的形成，而不仅仅是组分的线性组合。

（vi）一种重要的二级原理是对称破缺。如果系统的初始状态是完全同质的，进化着的结构便会过分对称。因为这将阻碍复杂结构的建立。对称破缺通常是藉遗漏或不正确地关联（或出现了其他的机遇），以及藉系统的非线性和所致的对小涨落的敏感性而自发地实现的。

（vii）另一个二级原理是入列（entrainment）。一些模式，在其受激发而行将登上舞台的意义上，可以捕获另一些模式[6]。这种过程增加了系统中的秩序，促进通过共鸣的联合的形成。

（viii）最后，也是最重要的原理，要求系统的记忆被分布式地加以存贮。记忆的重要性前面已经讲到了，在神经网络中，关联的强度即权重执行着这种存贮信息的功能。特定的权重不可能代表特定比特的符号信息，否则这意味着信息可以在权重的层次上加以解释。因为每一权重都只接近局域层次的活动，它不可能实施代表一个概念这样的更复杂的功能。复杂概念将包括涉及若干单元的活动模式。权重，作为记忆的痕迹。信息分布于许多单元的事实，不仅加强了系统的鲁棒性，而且使得不同模式的联合成为系统的内在特点——它们在原则上是交叠的。（第5章对分布式概念给予了详细的关注。）

这些基本原理使得自组织过程成为可能的方式，在后面特别是当讨论到大脑发展的“选择”理论时，我们将对此作进一步的阐述。下一节将讨论和论证这种观点：复杂系统不仅将组织其自身结构，而且倾向于以优化方式来实现。

6.3 自组织临界性

当我们面临着意外现象，特别是当它们具有灾难性后果时，我们倾向于将其归因于不太可能的情形的罕见组合。当我们必须解释股票市场的崩溃、地震的发生或是政治暴力的突然爆发时，我们力图发现导致它们发生的一些组合因素，常常还希望表明这种多因素同样组合再度出现的机会是微小的。然而，这种分析是试图将大的复杂系统的行为用小的简单系统行为的外推来加以解释的结果。不幸的是，这种外推是不成功的。复杂系统——其中许多因素以不同步的方式相互作用——展示了意外的且常常是不可预测的行为。任何分析，如果忽视了复杂系统的自组织行为的可能性，都将严重缺乏解释能力。

由巴克(Per Bak)、陈(Kan Chen)及其同事(Bak and Chen 1991)引入的自组织临界性概念，是分析复杂系统的一个十分有用的概念。这个概念，有助于我们理解由大量局域作用着的要素所构成的系统的某些全局性特征。

> ……许多合成的系统自然地向临界状态演化，其中微小的事件引发了能够对系统中任何数量的要素产生影响的链式反应。尽管合成系统产生出的是比突变要小得多的微小事件，但所有大小的链式反应都是整合的动力学的一部分。按照这种理论，导致微小事件的机制同时也是导致重大事件的机制。而且，合成系统绝不会达到平衡，而是从一种亚稳状态(meta-stable state)演化到另一种亚稳状态。
>
> 自组织临界性是一种整体论理论；全局特征，如大小事件的相对数量，并不取决于微观机制。结果是，系统的全局特征不可能通

过分别地分析部分而得到理解。据我们所知，自组织临界性是导致关于动力学系统的一种整体理论的唯一模型或数学描述。(Bak and Chen 1991：26)

一个简单的例子将澄清这个原理。如果你向一小块圆盘上倒沙子，将形成沙堆金字塔。沙堆金字塔的高度将主要取决于圆盘的大小以及沙子的特点。一旦大体达到了平均高度，沙子就会滚下来。但是，它是如何滚下来的呢？考虑一个已经达到了其“临界”高度的小沙峰，然后向上一粒一粒地加上沙子。这些沙子并不会一粒一粒地往下掉。它们可以停滞在沙峰上，或引起少许的雪崩现象。这个简单的模型中有两个方面是重要的。首先，一方面，沙峰将保持自身于其临界高度上下。当高度远不够时，更多的沙子将停滞于其上；当已有足够的高度时，更多的沙子将滚下来。其次，另一方面，任何个别的沙子都可能引起某种大小的雪崩现象：有时仅仅是若干沙粒，有时则是大滑坡。变化并不总是相同，有赖于沙峰是高于还是低于临界点，但是任何单颗沙粒的效应都是不可预测的。在此至关重要的是：系统向临界点方向组织其本身，在临界点单个的事件具有了最广泛的可能效应范围。换一个说法，系统调整自身以最优地敏感于外部输入。

一个常常用来形象化系统行为的方法，是在态空间对其进行描述。态空间对于系统的每一自变量都具有分别的维度。在 3 个变量的情形，比如说温度、体积和压力，态空间将是三维的。在 1000 个变量的情形(如同具有 1000 个节点的网络)下，态空间将是千维的。于是系统的所有的可能状态，将由态空间的唯一的点来标征，系统的动力学将形成通过态空间的轨道。当态空间中一些轨道奔向

某个点(或区域)时，该点(或区域)就是一个“吸引子”，代表了系统的一个稳定态。当轨道都在离开某个点，那个点就是不稳定的——称作“排斥子”。既有轨道奔向它又有轨道远离它的点，被称作“亚稳态”[7]。

节点的网络具有许多自由度，因此有一个大的态空间。将这样的态空间形象化是困难的或者是不可能的，但是此概念仍然是有用的。在非常稳定的系统中，可能只有一个或若干个强吸引子，系统会很快地达到稳定状态之一，并将不再轻易地移向其他状态。此种系统形成的行为不是很有趣。另一方面，在非常不稳定的系统中，没有强吸引子，系统将只是混沌地来回跳动。自组织临界性理论告诉我们：一个自组织系统将试图使自身平衡于僵硬和混沌之间的某个临界点。它将试图优化自己的吸引子数量而不会变得不稳定。为何这是重要的？显然，一个系统，如果仅仅是混沌的行为便是没用的。另一方面，一个系统，如果过分稳定则也是有缺陷的。如果每一所需的系统状态都必须是以强的稳定吸引子而出现的，那么大量的系统资源都将是受束缚的(在某点的自由度受限意味着许多节点必定是共享的)，系统的适应能力也将严重受损。而且，从一个稳定态运动到另一个稳定态也将需要非常强的扰动。因此，系统对于环境中变化的反应也将是迟缓的。不过，如果系统处于临界点，稳定态的数量将不仅受到优化，而且系统还将能够在最少努力的情形下改变其状态。

应该澄清的是，竞争原理是这种行为背后的推动力。网络中的每一节点都将倾向于支配态空间中尽可能大的部分，节点因此也为可利用的资源而竞争。向系统中的输入如果没有较多的多样性，将由少量的强吸引子来代表。随着输入的多样性的增加，系统将倾向于

临界点——在此灵活性得到最优化。如果系统必须处理的信息变得大于系统的内在能力，系统将被迫越过临界点。这将不可能产生任何稳定的吸引子，混沌出现了。由于这个原因，自组织系统的资源既不应过度，也不应不足。

可由这些现象来说明的自组织的另一个方面，是在仅有单元间局域相互作用的条件下大尺度特征的涌现。在这一方面，考夫曼(Kauffman 1991)提出了一些有趣的论据。他表明，在简单的网络中——每一节点要么开要么关，而且只有两种输入——是通过网络中形成的稳定性“核”(core)而达到有序的。这种稳定性然后通过网络而“过滤”，作为相邻节点而被已有的稳定群纳入稳定性之中。其结果是，网络中加入了“不变化的冻核和变化着的要素岛屿”(67页)。涨落群从而孤立起来，秩序加到了网络上。藉向节点加上某种“偏好”，这个模型可以做出更复杂的网络。这些将抑制一定阈值之下的信号——另一个非线性相互作用的例子。偏好的大小具有重要的效应：如果它太大，网络将过分稳定；如果它太小，网络将是混沌的。因此，偏好提供了一种机制，系统可以藉此调整自己，保持在临界水平，甚至当外部世界的复杂性起伏不定时也如此。

系统不得不朝向临界性的趋势，导致了复杂性的增长。如同考夫曼和鲍克这样的研究人员力图表明的，这种趋势是复杂系统的一种内在特征。一旦系统具有了自组织的能力，便会有某种“自然的”推动，以促使其组织优化。这种指向更复杂结构的推动力，是“经济”理由的结果：资源不能被浪费。在此方面要进行考察。系统的临界状态往往被叫做“混沌的边缘”(Lewin 1993)。“混沌”一词于是被在技术的意义上运用(决定论混沌，“混沌理论”的研究对象)，相应提出的论据认为，临界组织是某种有待混沌理论去发现的“深

层”原理的结果。认为自组织的临界性可以通过“经济”原理加以解释，而不必依赖任何来自混沌理论的论据则是一种既较为微弱也更为一般的主张。这并不必然意味着混沌理论是错误的。此主张仅仅是加强了这种论据(在前言中提出的)，“复杂性”是比“混沌”要更有普遍意义的范畴。在后面的各节中，当我们探讨自组织的一些进化的方面时还将强调这一点。

6.4 通过选择而组织

前一节主要关注的是复杂系统的内部结构。自组织是从一般原理上加以描述的，并重点研究系统的组分中的必然的内部关系，这允许有趣的动力学行为。在本节中，我将考察选择原理——此概念取自进化理论——应用于复杂系统中结构的发展。这里将再一次使用大脑作为中心例子。

复杂系统是开放系统，即与其环境相互作用的系统，不仅是能量上，而且是信息上。这些系统必须能够适应环境中的变化，因此它们的内部结构必定以某种方式受到外部条件的影响。通常所谓的系统“内部”与“外部”之间的截然区分，在这里变得有问题了。

系统在应对变化着的外部条件时的可能机制是什么？这里可以鉴明两种极端的情形。在一个极端，系统的结构完全是先验地定义的。这将意味着，系统是“硬连接”的，所有可能发生的事都必将由系统的固定的内部结构来驾驭。除了丧失适应性之外，这样的系统在复杂的情形下可能变得过分笨拙。在不那么复杂的情形下，以简单的控制原理来运行的“硬连接”系统，也许是一种适用的解决方案，但这对于我们感兴趣的复杂系统种类来说并非是一种合理的选择。在另一个极端，我们拥有完全没有独立的内部结构的系统，这

里的结构完全是由环境中的条件所决定的。一个系统，若只是直接地模仿环境，将不可能在那环境中发挥作用，因为它将完全受环境支配。要能够解释其环境，系统就必须具有至少如下两种属性：某种抵制变化的形式，以及某种比较不同条件的机制（用以决定在此是否有足够的变化以保证某种反应）。这两种属性仅仅转化为对某种形式的记忆的需求——不抵制变化，记忆就是不可能的。如果这种系统的自组织能力是合适的，它就将学习应付变化着的环境。

我们当然不局限于这两种极端。我们考虑生物系统时会受到一些约束。生命系统必须在极度复杂的条件下运行，因此适应性具有首位的重要性。在复杂的情形下，它还要受某种临时的约束。生物的生存，有赖于适当地及时作出反应。在某些关键条件下，或许没有充分的时间或余地(margin)对错误作出适当的反应，也许唯一的解决方案便是预定的“硬连接”反应（例如新生的哺乳动物就没有时间先“学习”如何使用其肠胃）。这表明对在两种极端之间找到某种形式的均衡的需要。但是这种均衡在何处？有多少是遗传上编码的以及有多少是可以归因于对环境的反应？在关于大脑的语境中，我们这里所得到的是关于自然与养成两难(nature/nurture dilemma)的再表述。

幸运的是，如果我们正好能够表明，实现系统的预决定需求和适应性需求是相同的机制，那么，精确地表明均衡所在也就变得不那么紧迫了。

6.4.1 通过选择而学习

进化理论试图解释的是生物系统如何通过世代更替而发展起某些能力，以提高生存能力。这种理论主要解释的是生物系统行为的预定方面，对任何特定的有机体的适应性则什么也没有真正说明。为

了保证其生存，一个有机体不仅必须学会如何应对变化着的环境，而且必须是在其自身的生命期间即所谓的“肉体时间”之内学会这种本领。如果我们把自组织的概念与进化联系起来，即如果我们能够发展起一种关于选择的更一般的理解，就可以认识到，在预定与适应行为之间的区分并非是那么界限分明的[8]。有两位神经生物学家，尚热(Jean-Pierre Changeaux)和埃德尔曼(Gerald Edelman)，在一定程度上已经尝试了这种综合。他们以稍微有些不同的方式将进化选择概念扩展到包括神经元网络中发现的适应性行为。埃德尔曼(1987)将之称为“神经达尔文主义”(Neural Darwinism)。

系统是如何对环境作出反应的？尚热等人(Changeaux *et al*. 1984)提到了两种类似于上面谈及的机制：

- 指令机制，环境将秩序直接施加于系统的结构上。
- 选择(达尔文)机制，在此秩序的增长是系统与环境之间相互作用的结果。环境并不决定系统的结构，但是影响系统中模式的发展，及其转化、强化和稳定性。

在神经网络的术语中，上面的区别可以用有指导的学习和无指导的学习来表述。这些术语的意义，在分析尚热和埃德尔曼两人都拒绝前一种选择的过程中就已清楚了。这个拒绝源于对这种思想的否认：世界在信息方式上是预排列的，即认为范畴是以某种先验方式划定的，而这些范畴是能够被客观地认识的。因此，这种拒绝，是对包括柏拉图主义和逻辑经验主义——形成了经典 AI 的理论框架——这类思想的拒绝。

神经系统的基本任务之一，是在一个"没有标签的"世界中进行适应性的感知范畴分类——其中宏观秩序和客体及事件的排列（以及甚至其定义或辨别）对于有机体都不是预定型的，尽管这样的客体和事件都服从于同样的物理规律。（Edelman 1987：7）

任何形式的学习——其中系统（这里是神经系统）直接地反映了世界——都意味着某种形式的指导：要么是由自然本身（这里必须提供范畴和它们的标签），要么是由一种雏形人（homunculus）。埃德尔曼直率地拒绝了这些直接的或"信息的"模型，特别是因为它们仍然形成了认知科学和人工智能中多数理论的支柱。因此值得以较长的篇幅引用他对于这些模型的分析：

按照信息处理模型，来自外围的神经信号以多种方式被编码，并随之被种种神经核和中间站进行转换；最后，它们由越来越精致的中继系统以种种方式再度转换，最终由大脑皮层进行处理并输出。必然地，这种观点非常强调大脑发展期间形成精确连接的严格规则。……这种观点也假设了，记忆的本性是事件通过记录或复制其信息细节的表征形成的。（38页）

于是应该清楚的是，"信息的"趋法严重地依赖于预定结构的概念——系统"自有的"能力，而不是系统可能获得的能力。而且，它假定了一种关于大脑的等级的基于规则的表征模型。这不仅仅是关于神经系统的内部工作的假设，而且事实上是更为基本的关于宇宙构造的本体论假设。它重新使得世界被还原成逻辑关系，从而可以被捕获。于是，对于复杂系统的建模，这种趋法与基于规则的趋法有

着密切的联系也就用不着奇怪了。而且，埃德尔曼意识到：

> 信息处理的概念，倾向于十分强调中枢神经系统计算物理世界相关变量的能力。这种观点的最终结果便是在假定大脑以算法方式进行计算的情况下进行算法和计算的讨论。……在物理世界中，自然对象的范畴被暗含地假定是落入到定义了的种类或类型之中，这是程序可以执行的。将此思想再推进一点，这种模型的某些翻版的支持者愿意这样考虑：在语言的句法结构和更高的语义功能的实现中出现的规则和表征(Chomsky 1980)，是在神经层次上相应的结构中产生的。如果统计变量最终进入这种观点下的大脑之中，就会被当成信号中的噪声来处理，而这在信息处理模型中被看作是变量的主要表现方式。(38 页)

埃德尔曼以“群体思维”(population thinking)来反对基于规则的或信息的观点，这是他的关于复杂系统的分布式本性的术语。下一节将更详细地考察他的趋法。

6.4.2 神经群团的选择

埃德尔曼(1987)希望解释大脑的较高级功能，或更特定一些地说是希望解释：如果没有“世界是以信息方式预排列的或大脑包含着雏形人”(4 页)的假设，感知分类是如何发生的。因此，他的论据不是基于大脑主要是由遗传预决定的结构构成的概念，而是基于大脑由大量简单的、未分化但是相互关联的神经元所构成的概念。这些神经元动态地组织成具有不同结构和功能的细胞群。组织作用主要发生在发育阶段，但这种组织变化在大脑生命期中的确都在进行。此过程依赖于三个原理(5 页)：

- 神经元的同质群体是通过一些选择性过程渐成地分化成结构多样的群团。一个群体是由若干神经元构成的，这些神经元具有强烈的与其他群团的关联性——由突触决定的关联性。这些群团形成了埃德尔曼所说的“基本技能”。
- 在其诞生之后，群团之中和之间的联结是通过作为系统与环境之间的相互作用的结果接收到的种种信号而渐成地变化的。群团的结构得到改进，更重要的是出现了与其他群团的相互作用（竞争与合作），导致了很可能在未来得到利用的功能群团的形成。它们形成了“次级技能”。
- 一旦基本技能和次级技能形成了，种种群团便会通过回路和反馈而相互作用。这种动力学相互作用是种种感知和原动反应的相关的结果，它引起了皮层映射的形成，使得大脑可以解释其环境并对之作用。大量的群团加入到映射形成中，特定的群团可以是许多个映射中的一部分。

我们可以用如下的一般术语对此模型进行小结。大脑以一般的、非特定的方式预先形成了结构（但是具有充分的分化，即充分的不对称），以允许外部的影响有“立足点”。于是，一般的“结构”通过经验和行为而变化，以反映所论的生物在历史进程中遇到的特定的情形。大脑从而这样组织自身来应对其环境。值得注意的是，基本技能的某些部分可以是永久性“硬连接”的，不会被经验所改变。这使得生物能够将一部分能力通过遗传传递给后代。这样的能力，会包括从一开始就必需的重要身体功能的控制，以及其他的生物所需的、但没有足够时间由自己来学习的能力。此种模型的美妙之处在于，硬连接和适应性成分都是以完全相同的方式运行着（即作为神经

元群），它们之间没有任何真正的区别。

在转向映射本身的形成过程之前，此模型的两个重要而又密切相关的特征需要加以讨论。对于第一个特征，埃德尔曼（1987）坚持认为，基本技能的神经群必须是“退化的”（6 页）。这意味着，在基本技能之中，有相当多的非一致变量群，其中每一变量群最终都可能与一定的输入结合起来。在基本群团中，没有什么以先验方式指定它负责某种特定功能。对于第二个特征，群团并非是封闭的单元，而是分布遍及大量交叠区域（163 页）。分布性概念是一个重要的概念，在埃德尔曼那里是采取“群体”趋法的直接结果。两个特征结合起来——退化性和分布性——否认了狭义上的皮层功能的局域化以及等级处理过程的存在（162 页）。

埃德尔曼讨论了皮层映射形成的三个重要的方面，阐明了映射形成的过程，并部分地结合了自组织的一般原理。这三个方面是群团限制（group confinement）、群团选择（group selection）和群团竞争（group competition）。

（i）群团限制。如果皮层只由同质的大量神经元构成，那么任何结构的发展都是不可能的。另一方面，神经元本身是不可能实施复杂功能的。神经群因此具有最优化的大小，但是不幸的是，它并非是恒定的或先验知道的。大小必须是自发地、动态地达到的。在皮层中，这是通过水平面上局域抑制的关联和垂直面上局域激发的关联来实现的。因为皮层是神经元薄片（皱叠地适合于头盖），活动被限制于垂直方向的群团，而与水平方向的相互作用较稀疏。这与已知的皮层结构是吻合的，可以作为大量密堆积“柱体”而很好地形象化（参

见 Mountcastle 1978)。

(ii) **群团选择**。群团选择主要是一种使用原理的功能(赫布规则)。群团被使用次数越多,就变得越是活跃。此过程主要是以如下方式工作的:承载从某种来源的(而无论是内部的还是外部的)脉冲的神经束激活某些神经群。当从一个以上来源的信息在某些神经群中交叠时,相关作用会导致那些群团中的高水平活动,在那些群团中的神经元将改变其突触的强度,变得对此相关更加敏感。在反复地暴露给一定的相关之后,此群团将对它作出强烈的反应,我们就说,这种相关"被选择"了。一旦被选择,此群团就变得活跃起来,甚至并非所有输入都同时存在时也如此。结果是,一定的联想得以形成并被记忆。如果相关只是非规则地出现或是不够强,突触的变更就将是不可能的,群团将再被其他过程的选择所利用。能够遗忘,是记忆的一个前提条件。

(iii) **群团竞争**。一旦大小适当的群团已经形成并被选择了,它们的稳定性就由其与其他群团的竞争相互作用所决定。微弱的小群团将被消灭,过大的或笨拙的群团将分裂并被较小的且更具生机的群团所征服,生机性(vitality)是由模拟和活动的频率与品质来决定。竞争对于"历史效应"也是非常敏感的(Edelman 1987: 165)。群团的功能是由其历史所决定的,现存的群团可以强烈地影响新群团的形成。通过群团的竞争,皮层映射得到了发展,并以稳定的但是动态的方式得以保持。不用的或不必要的群团被消灭了或"被清除了",不会削弱新群团的发展能力。

从这些原理中可以看到，映射形成不仅仅提供了关于大脑的认知发展的理论，而且还体现了所有自组织系统的三个基本特征：合作、竞争和记忆。

6.4.3 大脑与世界的关系

埃德尔曼的大脑模型的基本特征现在可以概括如下：

- 通过自组织过程在大脑网络中形成神经群团。
- 通过与环境的相互作用，这些群团被动态选择、变更和保持。
- （源自内部或外部的）脉冲引起某些群团中的活动，并通过在此过程中被激发的其他相关群团的活动，产生出适当的反应。

于是这里提出了一个问题：大脑以何种方式表征着必要的关于环境的信息？前一章的目的在于主张，经典的表征不是适当的机制，但是这里我们必须就环境与大脑结构之间的关系表述更多的东西。

指出环境直接影响着大脑的结构是重要的；它引起了大脑结构的变化。不过，在大脑结构只是作为其环境之镜的意义上，这些影响不是直接的。这里有两个原因：首先，效应是延迟的；外部影响效应的完全作用并不是立即被感受到的。因为系统必须有某种对变化的抵制，影响之效应的产生有一个过程，影响一旦开始以后就不会有最终的结束。第二个原因是，不同的脉冲总是混合的，这是因为群团是嵌入在网络之中的。这意味着，例如，尽管感知是一个直接的过程——网络中一定的脉冲是在直接反应中形成的——但它不会无中介地得以保持[9]。先前的脉冲（记忆）以及源于其他的脉冲，联合起来决定着反应的本性。这里所讨论的在神经系统与环境之间相互作用的本性，可以按照德里达的延异概念进行清楚的分析（参见第3章）。

大脑与其环境之间的相互作用的第二个方面，涉及原动行为的作用(Edelman 1987：8，45，209～239)。映射形成不仅仅是大脑过程的结果，因为可利用的信息也是由生物主动探索其环境的方式所决定的。大脑的结构受到行动的强烈影响。这支持了这样一种思想，智能并不是一种抽象过程，而是嵌入在世界之中的过程。我们的大脑是我们身体的一部分，反之亦然。

主动的原动行为的作用，形成了反对抽象的、唯我论的智能的前一半论据。其后一半涉及交流的作用。交流的重要性，尤其是符号系统(语言)的运用，不会使我们返回到客观的信息处理过程的范式。用于交流的结构也嵌入在神经结构中，因此将总是服从于网络相互作用的复杂性。我们的存在，既是具体的(embodied)，也是偶发的(contingent)。

6.5 自组织的哲学含义

自组织是大量真实系统的特殊特征，即是在我们周围世界中成功地运行着的系统。它可以从一般术语进行分析，并给予数学的描述。此种过程可以进一步建模，并在计算机中加以模拟，用于解决问题。从这种方式中获得的见解，对于我们理解这个世界及其居民，都具有重要的含义。本节将讨论一些哲学问题。

6.5.1 关系和模式的重要性

自组织是复杂系统的重要性质的主张，是反对基础主义(foundationalism)的。自组织的动力学本性，是不可能求助于某一个起源或某个永恒的原理来加以解释的，在此系统的结构由于偶发的、外部的因素以及历史的、内部的因素的相互作用而不断地发生着变化。事实上，自组织提供了这样的机制，藉此复杂结构可以进

化，而无需假定第一开端或超验干预。正是在此意义上，后现代理论对于我们理解复杂的、自组织的系统有所贡献。

由于类似于上面的理由，自组织系统也是反还原论的(anti-reductionistic)。作为相互作用的复杂模式的结果，系统的行为不可能只按照它的原子论组分来解释，尽管事实上系统除了由其基本组分及其相互关联组成之外，别无其他。复杂的特征是通过系统中的相互作用而“涌现”出来的。加芬克尔(Garfinkel，1987：202～203)在生物学语境中讨论了部分和整体之间的关系：

> 我们已经明白，对聚集体进行建模需要超越个体细胞的水平，以整体的变量来描述系统。但是，在经典的还原论中，整体实体的行为最终必须以参照其组分的本性来加以解释，因为这些实体“正好是”低层次对象及其相互作用的集合。尽管在某种意义上它可能是正确的，系统“正好是”其要素的集合，但是这并不意味着我们能够参照其部分并结合关于它们的关联的理论来解释系统的行为。特别是在处理由大量相似组分组成的系统时，我们必须要使用将系统行为看作整体的整体性概念。我们在这里遇上了这样一些概念，诸如入列、全局吸引子、聚集波。尽管这些系统的性质最终必定是可以按照个体状态来定义的，但是事实并不使得它们是“虚构的”；它们是因果有效的(因而是真实的)，并与其他的系统变量有一定的因果关系，甚至对于个体的状态也是如此。

自组织的相关性在接纳某种尝试联合和包括，而不是诋毁和忽视的系统思维(systems thinking)时就变得清晰了。这是一种不是被矛盾和对立所吓倒，而是转向这种赋予系统活力的力量的思维。自从

哲学出现以来，形形色色的系统思维是相伴而行的。赫拉克利特(Heraclitus)的思想是一个好的例子。对于他来说，宇宙的基本原理是斗争：战争对所有的事物都是共同的，斗争就是正义，万物都是通过斗争而出现，通过斗争而消亡的。不是偏惠于某种特定元素——泰勒斯(Thales)的水、阿那克西米尼(Anaximenes)的气——赫拉克利特把万物置于互相的竞争之中。在这种动力学张力中，"万物都出现并消亡"。系统思维也支配着中世纪的四元素理论及其各自的特性和原理(Flanagan 1990：80～105；Wilden 1987：151～166)。当所有的元素都处于平衡时，当它们都合作时，事物的理想状态就出现了。人们相信，在世界的堕落状态下是不可能达到这种和谐的，因此就需要进行不断的调整。系统思维的现代例子，会包括索绪尔的语言学和弗洛伊德的心理学。后现代否认单个的元叙事，强调差异和对立的重要性，虽不是常规地以"群体思维"(这里是运用生物学表达的系统思维)术语来阐述，但是其相似性是不容否定的。

尽管珍视关系之重要性的思路——以及追求模式而不是本质——在整个的西方智力史上都是随处可见的，但是这通常都被更为勇气十足地宣称已发现真理的理论主张所压抑了，这种理论如：柏拉图理念论、理性主义、马克思主义、实证主义。在我们对于复杂系统(如大脑和语言)的分析中，我们必须避免落入找到万能钥匙的陷阱。由于复杂系统建构自身的机制，单个原理提供不了适当的描述。我们反倒应该对于复杂和自组织的相互作用保持敏感，并珍视模式的作用，它们不停地转化着系统自身以及它们所置身的环境。

6.5.2 历史的作用

记忆在自组织系统中起着重要的作用，这已经被反复地讲到了，但这个主题还是值得在哲学语境中再进行一下扼要的讨论[10]。

复杂系统，无论是生物系统还是社会系统，如不考虑其历史，就是无法理解的。两个类似的系统，如果具有不同的历史，则即使将其置于一致的条件下其反应的方式也可能有极大差异。更准确地说，系统的历史不仅仅在对于系统的理解中是重要的，它还共同地决定着系统的结构。

然而，在分析一个系统时，应该谨慎使用历史的概念。它不应被看作是开启了系统的最终描述之门的一把金钥匙。这是因为，一个系统的历史，不是以能够被重构的方式在系统中呈现出来的。系统历史的“效应”是重要的，但是历史本身是通过系统中的自组织过程而连续地转化着——留下的只是分布于系统中的历史痕迹。这些痕迹，在神经网络中是藉相互作用及其权重进行建模的，并不相当于事实、思想或符号，因为它们是在极其低层上对信息进行编码的。系统的全局行为是“痕迹的模式”的结果——构成模式的单个痕迹其自身是没有意义的。单个痕迹可以是属于一个以上的模式，而同样的模式并不总是由同样的痕迹集合所构成。因为痕迹处于不断地转化之中，严格说来，是不可能精确地再次得到“同样的”模式的。任何复杂的动力学系统都由于其环境和本身两方面而处于不断地转化之中。

这样的论据，对于大脑的语境中的记忆也是适用的。记忆不是作为离散单元存贮在大脑中，如同可以从档案柜中取出一样。记忆痕迹以分布方式存贮，并由于经验而不断地转变着。在言说方式中，尽管记忆是所有大脑的高级功能的基质，但却没有“复数的记忆”(memories)。这如同“分布式表征”的概念。信息“涂抹”遍及许多单元的事实，是复杂系统的一个至关重要的特点，而不仅仅是对于图标式表征的一种替代。

6.5.3 主动和被动

自组织的另一个哲学上重要的方面，是过程有些滞置于(suspended)主动和被动模式之间[11]。一个自组织系统，会反作用于环境中的事态，同时也作为这些事态的结果而转变自身，而常常又反过来对环境产生了影响。系统中的过程因此既非对外部的简单的被动反映，也非被内部主动地决定着。在主动和被动之间以及内部和外部之间的这种区分，成为了被迫而为之的事。在一个复杂系统中，控制并非由单个来源所致。如果是这样，系统将会退化，失去其适应性，只有在环境保持稳定时才能生存。这些概念，对于我们思考社会系统是有意义的，这将在本章的结尾处加以讨论。

自组织的另一个与主动与被动之间的滞置密切相关的方面，是过程的自反性(reflexivity)。自组织是一种自转化(self-transforming)过程；系统对自身发生着作用。自反性不允许对系统的任何静态描述，因为不可能半途截取自反时刻。自反性也不允许在元层次上对系统的完整描述(Lawson 1985：20，21)。一个元层次总是可以建构的，但是从这种观点来看，这只有在系统存在的给定时刻的瞬间才是可能的，即过去的遗迹与未来的预兆不发生结合的冻结时刻才是可能的。由自组织系统的自反本性所产生的时间的复杂性，是不可能被呈现出来的[12]。

我们对于复杂系统的理解，不可能忽视自反性的作用或滞置于主动和被动之间的运行模式。不过，当试图对复杂系统进行建模时，这些方面的确导致了严重的困难。形式模型目前的局限，并不排除建构出把复杂系统的显著特点结合起来的“机器”的可能性，但是应该记住的是，这样的机器将与其模型同等复杂，因此是同样地难以进行分析的。一个复杂的神经网络可以作为这种“机器”的一个例子。

6.5.4 稳定性和预见性

在对自组织的临界性的讨论中，有序和无序之间的精巧均衡的必然性得到了强调。当分析系统的稳定性时，或当试图预见其行为时，进行的也是类似的考虑。

经典的稳定性定义认为，在一个稳定系统中，小的原因产生出小的结果。在临界地组织起来的系统中，情况不是这样的，而从经典考虑的角度看它应该被称作是不稳定的。不幸的是，只要涉及活的复杂系统，经典的稳定性便会被归结为停滞和死亡。对“稳定性”概念的重新解释从而是必要的。

经典的不稳定性的定义，至少在庞加莱(Poincaré)的使用中，是概率的定义。不稳定事件被定义成不具备可观测原因的事件(Pattee 1987：328)。它们因此是偶然事件(chance event)，与决定论事件相反。不过，在复杂系统中，新奇的不可预测行为不必是偶然的结果。它可以由大量因素的复杂相互作用而“引起”，这些因素还可能是逻辑上不相容的。复杂性不可以与随机性(randomness)和偶然性混淆起来，但又不可能以一阶逻辑术语进行描述。“我发现没有其他的办法，而只能接受多重的、形式上不相容的描述，作为对于许多类型的生物学事件的满意解释”(Pattee 1987：329)。正是复杂约束的相互作用产生出有趣的行为——这些行为不可能被描述成偶然的或不稳定的事件。一个基于偶然事件的理论，包括量子理论，“只是作为从经典决定论的逃离：它不是自组织理论”(Pattee 1987：330)。

这将我们引向了预见性问题。能预见一个系统的行为，是一个成功理论的经典标准之一。然而，我们在与不可能完全以经典理论来描述的系统打交道。预见它们的行为可能是非常成问题的。克罗霍(Krohn)和屈佩尔斯(Küppers)(1989：155，156)对

这个问题的描述是：

> 这些"复杂系统"(Nicolis and Prigogine 1989)或"非平凡机器"(von Foerster 1985)，对输入 —输出关联的功能分析必须补充对其"机制"的研究，即补充因果分析。复杂系统，由于其操作条件，如果不考虑它所产生的机制，那么其输出(按照功能的或者预期的结果)就几乎是难以理解的。系统的输出联系着系统的"历史"，其本身有赖于将先前的输出作为输入(操作上的封闭)。系统的发展是由其机制所决定的，但却是不可能预见的，因为在其输出本身不可能找到可靠的规则。甚至更为复合的系统，其中起作用的机制本身也是按照归复式操作而发展的(学习的学习，发明的发明，如此等等)。

借助于对系统固有的主要约束因素的勾画，以及关于系统的历史和环境的某些知识，预见是可以进行尝试的，但从来都是没有把握的。这样的考虑，对于许多被迫应付复杂环境的社会系统具有重要实际意义。例子中将包括：动荡时期中的政治团体，被迫为有限基金而竞争的研究工作，在起伏的经济中的商业组织。因为未来可以预见的确定性已经大大地减少了，所以任何的行动计划都必须不断地适应。如果计划过于僵硬——过强的中心控制——那么系统将不可能应付不可预见的变化。另一方面，如果对所有的表面性变化都努力地调节本身，那么也将是灾难性的，因为这样的变化可以轻易地不知不觉地倒转过来。如果系统试图追踪所有的涨落，而不是适应其更高的有序趋势，那么便会浪费资源。能够将那些应该追随的变化与那些应该抵制的变化区分开来，对于任何组织(或生物)的生存都是

至关重要的。当对于系统的控制不是僵硬的、局部的，而是分布遍及系统的，保证了正面的自组织动力学被有效地利用，就有可能优化地做到这一点。

最后，我们将考虑自组织过程对于伦理学的启发含义。

6.5.5 自组织的伦理学

传统的对于进化论的批评，以及扩展到对于自组织的批评，在于它以过分的机械论看待人类的存在。这种批评指出，人类不可能仅仅从利牙坚爪的存在来加以理解。我们的存在也是一个由价值支配的存在，价值赋予生命以意义。这样的批评之所以可能，当然是以某种形而上学承诺为基础——或是以宗教为基础，或是以某种普遍的、抽象的(康德式的)原理为基础——它总是密切地联系着诸如不存在后现代伦理学这样的主张。

自组织理论是否主张有可能不借助某种一般价值系统来描述人的行为？如果这个问题中的价值指的是与我们的偶发性存在相分离的价值或者优先于我们的偶发性存在的价值，那么答案必定为“是”。然而，对价值还有另外的一种理解，它不仅与自组织理论是相容的，而且可以被看作是它的结果。按照这种观点，价值被理解为社会系统的涌现性质。我将给出两个例子：

- 对进化(或自组织)系统的建模已经表明，系统成员的纯粹自私的行为不仅对于系统是有害的，而且最终也将导致对于特定个体是有害的。因此，利他行为就不只是由“美德”个体采取的“价值”；它是对于系统的生存和繁荣的特征性的必然(参见 Axelrod 1984)。
- 我们已经看到了，分布式的、去中心化的控制使得系统更加灵

活，因此增加了它的生存性。如果我们将这个概念运用于社会系统，便有了论据来反对诸如公司管理或国家运行中的僵硬的、中心化的控制机制。而且再一次地，这种批评也是反对独裁管理或法西斯政府的，这不仅是以思想——这些事物本质上是“恶德”——为基础，而且是以知识——它们将最终导致所论系统的退化——为基础的。

这些性质意味着某种自组织的伦理学吗？我在这里是非常犹豫使用“伦理学”一词的，除非在这样的意义上，即它是一种“构成系统的原理”，接近于列维纳斯(Levinas)使用这个词的方式。这显然距离公共道德规范还差得相当远。这些概念的意义将在最后一章中更详细地加以探讨。

注　释：

1. 需求的结构性与需求的可塑性之间的张力正在不断增加。某种形式的结构对于系统中的系统编码是一个前提，但是如果结构过于僵硬，它就难以适应环境。类似地，可塑性对于适应性是必要的，但是如果变化的发生过于容易——如果系统中的记忆过于短暂——那么系统就只能反映其周围，而不能对它加以解释。这将在后面受到更多的关注。

2. 元胞自动机(Serra and Zanarini 1990：50～60；Toffoli and Margolus 1987)以及数字元胞自动机——也被称为随机布尔网络(Serra and Zanarini 1990：64～69；Staufer 1987)——提供了另外的模型。布尔网络也被考夫曼运用于他的自组织系统的模拟。尽管考夫曼无疑做了重要的工作，我将不对他的工作进行讨论。问题部分在于，对于从哲学层次进行讨论，难以找到考夫曼的有关文本。如果他写一本书，平衡于严格数学的《有序的起源》(Origins of Order)(Kauffman 1993)与花哨而准宗教般修辞的《在宇宙的家中》(At Home in the Universe)(Kauffman 1995)，那将是富有帮助的。

3. 在大脑(它是神经网络)的情形中，接收到的刺激不仅仅是通过感觉而来自外部，而且也来自身体的内部。

4. 这些原理是以 Von der Malzburg (1987：272)提供的类似清单为基础的。

5. 预先定义的结构并没有被排除掉。在活系统中，重要的结构是可以继承而来的。

6. 在自然中发现的人列的例子包括：蟋蟀互相接起来成为相继连贯而鸣的群体；萤火虫运动成为连贯发光的群体；密切生活在一起的妇女其排卵周期趋同(Garfinkel 1987：200)。

7. 摆锤静止在竖直向上的位置的钟摆是亚稳态状态最好的例子，最小的干扰立即会使它离开平衡状态。

8. 理想上我们需要以公式表示选择原理，这应与赫布规则(或者使用原理)是相容的。

9. 参见 Edelman (1987：234，235)关于直接感知理论(Gibson 1979)与神经元群选择之间的差异的讨论。

10. 结合了弗洛伊德和埃德尔曼观点的关于记忆作用的一个重要讨论参见 Rosenfeld (1988)。罗森菲尔德主张非符号、非表征的记忆理论。

11. 这种洞见要归功于高夫斯(Andries Gouws)。

12. 参见 Cornell (1992：116～154)关于复杂系统语境中差异动力学如何解构时间的线性理解的讨论——在康奈尔的案例中是法律系统。

第 7 章　复杂性与后现代主义

7.1　承认复杂性

无论我们是否乐意将我们生活的时代叫做“后现代”，不可否认的是，我们生活的世界是一个复杂的世界，如果我们要生存甚至是要取得成功，就必须面对这种复杂性。传统的(或现代的)面对复杂性的方式，是试图找到某个参照点作为根基，从而越过断裂，成为一种所有事物都可以从中派生出来的万能答案。无论这个参照点是什么——是理想观念的超验世界，是激进怀疑论的精神，还是现象学的主体——在我看来，那都是构成了力图避免复杂性的策略。去找到本质真理的欲望，使得我们无法看见复杂性的关系本性，特别是那些关系的不断变化。任何对复杂性的承认，都必须将这些变迁和变化结合进来，不是作为附带现象，而是作为复杂系统的组分。

在这里的研讨中，我们到目前为止所关注的是一般性技术水平上关于复杂性理解的进展。在这个过程中，我们不断地比较相反的两种趋法——形式的、基于规则的、表征的趋法，以及分布式的、自

组织的、联结论的趋法。已经得到的一般性结论是，基于规则的趋法对于复杂系统的一般特点不具有充分的敏感性。已经指出的是，联结论模型要好一些，因为它们对于复杂系统的关系特点是内在敏感的，这也许是因为其基础是复杂系统中一个我们已知的最好样本——大脑——的事实。甚至尽管在目前，它们的能力与哺乳动物的大脑相比都还相差很远，但是它们至少是结合了复杂系统的一些一般特点。

在这一章中，我希望指出的是，对复杂性的承认有助于我们面对一些重要的哲学问题，包括伦理学和政治学的哲学问题。我将分析复杂性关系模型与后现代主义之间的关系。我们将提出一种论据来反对这样的观点：后现代趋法意味着“怎样都行”。相反，我们的看法是，后现代趋法对于复杂性是内在敏感的，这是对自组织的重要性的承认，对传统表征理论的否认。这里将聚焦于一种常被引证的后现代理论：利奥塔尔的《后现代状况》(The Postmodern Condition)。我还将考察，这种复杂性特性对于我们理解语言、社会理论、知识主张的状况以及伦理学的一些含义。

7.2 后现代主义

“后现代”一词拥有如此多种的不同意义，以致要对它下定义变得不可能了。关于后现代主义的文献已经激增到如此的程度，以致要准确地知道其中的某人究竟是在干什么也是困难的。有时，“后现代”和“后结构”这样的术语是合并使用的，而在其他时候，它们之间又被划出了清楚的界限；有时，“后现代”被用作某种严格的理论术语，而在其他时候，它又被当作某种模糊的术语用来描述当代的文化景象。为了能够有某种聚焦，我将关注在此领域中的一个经典

性文本——《后现代状况》(Lyotard 1984)。作出这种选择的推动力，是我在这里感兴趣的现代知识的状况，而不是进行一般的文化考察。我并没有要向后现代表示某种道歉的意思，也不想作出某种相应的解释，而是希望从我们的复杂性概念模型，对当代社会和社会理论进行分析。我将提出，后现代理论中描述的话语和意义的扩散，并不是由任性的、破坏性的理论家创造出来的，而是我们的语言和社会空间的复杂性之不可避免的后果。信息激增，以及媒体将国际公共空间塞进局部私人空间的方式，阻止了我们形成对我们的世界的统一的、连贯的描述。

从这里出发，利奥塔尔建立了他的后现代主义的描述。其目的是要研究发达社会中知识的状况(xxiii 页)。他认为，科学知识的习惯是诉诸连贯的元话语而使自身合法的，元话语起着某种一般性统一的功能。如能找到这样一种元话语，便可能将所有形式的知识都结合成一个宏大叙事(grand narrative)。这就是现代主义的梦想：

> 我将运用现代这个术语来指任何一种科学，它以明确地诉诸某种宏大叙事的元话语为参照来使得自身合法，诸如精神的辩证法、意义的解释学、理性或工作主体的解放或财富的创造就是这种元话语。(xxiii 页)

后现代主义于是被定义成“对元叙事的怀疑”(xxiv 页)。我们不是要找寻可以将所有形式知识统一起来的某种简单话语，而是必须处理话语的多元性，即许多不同的语言策略——它们都是局域地决定的，而不是由外部合法化的。不同的制度和不同的语境会产生出不同的叙事，而不同的叙事无法相互归结。

> 叙事功能正在丢失它的功能者、它的伟大的英雄、它的伟大的旅程、它的伟大的目标。它已经被散布于叙事语言要素的云团之中——叙事是概述性、说明性和描述性的，等等。每一云团中的传递，都有相应的实际效价。我们每个人，都生活在许多这样的岔口。然而，我们的确不必建立稳定的语言组合，我们的确已建立的这种组合也不必是可交流的。（xxiv 页）

有必要强调指出，这种将知识描述为局域叙事多重性后果的主张，并非反对科学知识，而是反对对此种知识的某种理解。利奥塔尔拒斥的是，将科学解释为代表所有真知识的总体。他认为，为了叙事地理解知识，将其描绘成多元性的较小故事，可使它们在其使用的特定语境中很好地发挥作用（7 页）。不是主张知识的不可能性，而是“它改进了我们对于差异的敏感性，并增进了我们对不可通约的宽容。其原理不是专家的异体同形（homology），而是发明者的形似性（paralogy）”（xxv 页）[1]。

让我们总结一下利奥塔尔的见解。不同的群体（机构、学科、社区）讲着关于他们所知所为的不同故事。他们的知识并不采取逻辑结构和完整整体的形式，而采取叙事的形式，这有助于允许他们去实现其目标，并使得其所为具有意义。因为这些叙事都是局域的，不可能连接起来形成某个将所有的知识统一起来的宏大叙事。后现代状况是以异质话语的多重性共存为特征的——一种不同部分对事物有不同评价的状态。那些对于统一的元叙事充满着怀旧情绪的人们——作为在西方形而上学史上占据中心的美梦——经历着把后现代状况看作是破碎的，充满着政治混乱，因此最终是无意义的。这使得他们感到晕头转向。另一方面，那些拥抱后现代主义的人发现了它的挑

战性、它的激动人心，以及广阔的空间。这使得他们感到了冒险的味道。究竟采取两者之中的哪一种评价，这常常取决于人们对哪一种感到舒服，而没有固定的参照点。在两者之中的选取，如同理论考虑中的情形，是由心理学决定的。不过，在这里的研讨中，我将只考虑理论问题。

有一个重要的、常常用来作为反对后现代主义的论点（参见如 Parushnikova 1992）。这种论点主张，如果所有的叙事都只具有局域合法性，所导致的社会之网的破碎将会使得所有知识相对化。因为不存在对于任何话语的外部“检查”，任何局域叙事都不可能受到严肃批判。每一种话语都将独立于所有其他话语，这导致了散漫的共同体的封闭和孤立。最终这将意味着，每一个体都只将自己作为一个参照点，完全没有任何东西可以奠定知识的客观性。其后果将是“怎样都行”——一种显然不可接受的见解，尤其是在科学哲学的语境中。

反对后现代主义的一般论点——对于固定参照点的否认将意味着怎样都行——在我看来，至少是令人失望的。它等于说，如果太阳不再发光，那必定是黑暗。我希望指出，至少就利奥塔尔所关心的而言，后现代主义不必承担“怎样都行”的见解。“怎样都行”的论点，取决于对于个体作用的某种理解，而这是利奥塔尔明确拒斥的（1984：15）[2]。

> 对于宏大叙事的打破……导致了某些作者从这样的角度荒谬地进行分析：社会组带的解散和社会聚集的解体，致使大量的个体原子被扔进布朗运动。绝没有这种事情发生：这种观点，在我看来，是被失落的“有机”社会的天堂似的表征缠住了。

对“怎样都行”见解的支持者而言，拒斥将个体理解为“有机”整体中作为“自然”角色的孤立的、自主的行为者，与接受将个体视作破碎的、原子论的观点是同义词。对利奥塔尔的仔细研究表明，他是在对于宏大叙事的消解所形成的破碎和孤立的思想加以拒斥的立场上来理解个体的。如下的引文包括了将被仔细分析的重要观点。它还阐明了，利奥塔尔的后现代社会模型与本研究中(包括自组织的特点和分布式表征)发展起来的联结论模型之间的关系。

> 一个自我并非分量很重，但是没有一个自我是孤岛；每一自我都存在于前所未有的更复杂、更多变的关系交织之中。不论是年少还是年长，男人还是女人，穷人还是富人，一个人总是定位于特定交流圈子的“节点”上，无论这些是多么微小。或这样说更佳：一个人总是定位在某个位置上，种种的消息都由此通过。没有任何人，甚至是我们之中的最无优越性的人，面对横穿并将之定位成发送者、收件人或指示物时，都不会是全然无力量的。与这些语言策略效应（当然，语言策略都是有关的）有关的流动都是可以宽容的，至少在一定的限度之内是这样（且此限度是模糊的）；甚至为规则的机制所诱惑，以及特别是系统为了改进其性能所采取的自调整。甚至可以说，系统能够并必须大力鼓励这样的运动以便与其自己的熵战斗；某种意外“移动”的新奇性，与其伙伴或伙伴群体的相关位移，能够从它不断需求和耗费的性能之增加而得到补充。（15 页）

与联结论模型的相关性，在这段话中得到了清楚说明。自我以“关系交织”来理解，是网络中的一个节点，而不是以原子论单元来

代表它们自己。因为重要的是关系，而非这样的节点，“一个自我并非分量很重”。利奥塔尔关于后现代状况的描述，事实上是对于我们社会的网络的描述，是对处于其中产生出并再生产知识的描述。他的要点在于，这种网络已经变得太复杂，难以进行一般性的或包罗万象的描述。复杂系统的所有特点(如第1章中所描述的)都可以在其中找到——一些正是现在要讨论的。关于话语多重性的论据，不是某种任性的运动；它是对于复杂性的承认。它允许信息爆炸以及作为真正复杂网络部分的不可避免的矛盾。

认为后现代的见解会产生许多孤立的、怎样都行的话语，在此可以从两方面加以反驳。首先，社会形成了网络。尽管不同的话语形成了网络中的不同“簇群”，但它们不可能将自己从网络中孤立起来。它们总是与其他话语相关联。不同的局域叙事是相互作用的，可以是一些比另一些更多一点，但是任何话语都不是固定不变的或是由自己来稳定的。不同的话语——网络中的簇群——可以生长、收缩、分裂、接合、吸取其他话语或被其他话语吸收。局域的叙事，只有与周围叙事有对照和差异才是有意义的。我们所具有的，是一种自组织系统，其中意义是通过动力学过程产生出来的，而不是通过“怎样都行”的自治性行为者的被动反映产生的。话语不是处于自足和孤立状态，而是处于不断的相互作用之中，在为了领地而相互斗争，它们之间的暂时的边界是游戏之中的划界。这就是利奥塔尔所称的“社会的竞技方面”(16～17页)。

关于社会网络的第二个方面，网络的分布式本性是与关于孤立作用的思想相抵触的。在一个联结论网络中，信息不是由特定的“节点”所代表的，而是在分布遍及于许多节点的模式中被编码。反过来，任何特定的节点形成了许多不同模式的部分。在社会网络中，

话语也是类似地散布于许多“自我”。 话语仅仅是一种交换着局域信息的遍及大量个体的群团的“活动模式”，而且并不代表元叙事的任何方面。 每一个体也是许多模式的部分。 一个人，可以同时是母亲、科学家、消费者、政治活动家、艺术家和情人。 因为社会网络共享着联结论网络所具有分布性的特点，所以，认为后现代主义导致孤立的见解就不攻自破了。

利奥塔尔涉及社会网络的进一步的方面，在上面引用的段落中是关于自组织——“系统为了改进其性能所进行的自调整”。 我已经指出这样的事实——放弃决定论的元叙事是一方面，而放弃自治性行为者是另一方面——将自组织过程滞置于主动和被动模式之间的某处。社会网络的动力学，也共享着在先前章节中所讨论的自组织的其他特点。 社会之网不是根据某种超验的原理来“设计的”，而是在以动力学方式对偶发信息的反应方式中形成并发展起来的。 此过程是复杂的过程，涉及许多个体，它们之间具有复杂的非线性的关系，包括反馈关系。 个体结合起来形成了簇群，但是仍然在网络中为资源而竞争。 因此，系统不是且永远不会是对称的(这一点我们随后将讨论)。 对于在其中任何部分的意义产生方式，系统的历史都是极其重要的。 社会之网中引起了连续的变更的结构进化，是其动力学一个整合的部分。

利奥塔尔对此是相当清楚的，社会系统的复杂性并不自动地导致随机性或噪声。 从上面的段落看，业已清楚的是，系统“与熵而战”，产生的是意义，而非噪声或混沌。 为优化此过程，系统必须尽可能多样化，而不是尽量地有结构。 在利奥塔尔看来，这是形似性功能。 形似性因此发挥着类似于自组织临界性(前一章中进行了讨论)的功能。

自组织的临界性是网络尽可能多样化其内部结构的机制。结构越是多样化，可以存贮和操作的信息就越是丰富。网络必须走在以固化的结构为一方和以无序为另一方之间的钢丝绳上。在我们的网络模型中，这个过程是单元或单元群团之间的激烈竞争的结果。在利奥塔尔看来，社会系统中的推动力是形似性和意见纷争(dissension)：

> ……在此有必要对意见纷争加以强调。共识是一种永远也达不到的地平线。研究发生在范式的庇护之下，倾向于稳定化；这如同技术、经济或艺术“创意”的利用。它不能被打折扣。但是，令人印象深刻之处在于，有人总是要扰动“理性”的秩序的。有必要假定存在着使得解释能力不稳定的力量，这种力量展现在新颁布的理解规范中，如有人乐意，也可以展现在为新的研究领域的科学语言建立新规则的建议中。(61 页)

利奥塔尔坚持把意见纷争和不稳定的力量作为共识的反面，因此概念也成为利奥塔尔批评哈贝马斯(Habermas)的核心，这对于整个哲学尤其是科学哲学，具有严肃的含义。科学的作用，传统上一直被理解为必须将知识固定在永久的网格上。实验证据被用来证明理论。充分地确证将保证其在网格上的永久位置。然而，很快就清楚了的是，客观确证的条件是有问题的，实验证据可以支持一个理论，但是却无法证明它。实验的过程不可能包括所有可能牵涉的因素，也不可能预见新知识如何改变对实验结果的解释。因为人们还可能否证理论，确证的过程就被证伪的过程取代了。如果人们不可能向网格上添加新东西，那就成了不够格的多余的一员。这种“扔掉”

策略，其结果是使得知识主干有资格作为“科学的”(scientific)部分越来越瘦弱。所有过于复杂的或包含着不确定性或不可预见性的事物，至少在暂时都要被搁置一旁。结果是，总体的人类知识的大部分，被轻视成不科学的(unscientific)——大多数的艺术、大多数的心理学(对许多科学家来说，弗洛伊德仍然是处于科学江湖骗子的范式中)，而且常常更一般地，人文科学(human sciences)也属非科学之列。以这种狭窄的关于科学是什么的理解来工作，使得甚至生命科学(生物学)以及经验科学(工程学)都成为可疑的。将其推向极端，证伪理论意味着，只有抽象的、先验的真理才是真正科学的。

利奥塔尔提出，我们忽视了共识的思想，是因为它处在贫困中。为了增进知识，我们必须增进话语，而无需试图将其固定在永久的网格上。这种见解与费耶阿本德(Paul Feyerabend)(1975)的见解有一些密切联系。费耶阿本德坚持一种科学的“无政府主义”，其中所有的边缘声音都应该参加进来。在此没有永恒不变的“方法”——正是它们决定着什么形成准绳部分，什么不可能。人们不应该将所有不符合此框架的东西都扔掉，而应该致力于发现在不同话语中的富有意义的关系。

在这方面，联结论模型提供了一种极其重要的洞见。如果承认所有知识都是嵌入在更大的社会网络之中——对这种观点的接受仍然是一个分水岭——意义和话语的增进就是复杂自组织网络的一个不可避免的特点。利奥塔尔和费耶阿本德并不是任性的反科学的(anti-scientific)无政府主义者；他们在考虑复杂社会中知识的状况。为使得先前边缘的声音再度获得平等的机会，并不意味着“怎样都行”。纷争之声并没有获得任何的优先权；它们必须进入到“网络的竞技场”，其中它们的相关性，是根据系统的历史以及变化着的需

求和目标，通过竞争和合作而动态地被决定的。

这里以一种谨慎的评注作为本节的小结。因为我们所谈到的所有网络都是条件性实体（contingent entities），它们是有限的。甚至最复杂的网络也只是具有有限的处理信息的能力。因此，一个网络将难以承受过度负担，当面对着过多新奇性时尤其如此。一个过度负担的网络，将显示出“病态”行为，或是混沌行为或是紧急关闭。这实际可能成为许多后现代主义批评所害怕的事态，其中之一是我们处在过度的信息负担之中，并淹没在此过程之中（如 Baudrillard 1988）。然而，要点在于，很少有可能幸免。回复到网格，采取中心控制或重新引入宏大叙事，但都不会使得信息消失。我们必须学会处理它，通过更有效的区别对待，过滤掉某些多余。

而且再一次，联结论模型是实施这种“过滤”最有效的模型。在基于规则的系统中，偏好必须被编入程序，调整也是困难重重。这样的系统仍然属于范例性的现代主义的趋法，即以抽象形式的意义（表征）以及中心控制来工作。联结论模型可以动态地调整自身，从而选择何者应该被抑制，而何者应该被增强。鲁棒性和灵活性是同一枚硬币的两面。在我们的社会状况下，这意味着，如果我们之间以及我们与环境之间的相互作用变得不那么僵硬，那么我们将经历较少的后现代压力。这并不意味着人们只好放弃，或者随波逐流。而是意味着，我们所有的人，都不得不进入到网络竞技场之中。

7.3 复杂系统与后现代社会

在本节中，我希望主张，后现代社会（看作一个系统）能够从第 1 章中描述的复杂系统的 10 个特点的角度来加以描述。对此我将逐一讨论，并指出其中一些对于当代的文化理论具有的有趣意义。

（i）**复杂系统由大量的要素构成**。如果我们把社会系统作为由人的个体构成的，那么系统中要素的数目肯定是巨大的。

（ii）**复杂系统中的要素是动态地相互作用着**。个体处于不断的信息交换之中。记住，在神经网络中的特定节点其意义有限，是相互作用的模式在编码信息并产生出意义。类似地，富有意义的不是人的个体的孤立存在："一个自我并非分量很重"(Lyotard 1984：15)。个体是由对于他者的关系构成的。

（iii）**相互作用的层次相当丰富**。人的个体在一个极其巨大的差异能力的队列中与其他许多个体相互作用。在后现代社会，相互作用的层次在不断地增加。

（iv）**相互作用是非线性的**。非线性是复杂性的先决条件，特别是自组织、动态适应和进化关系重大时。与非线性原理密切相联系的是不对称性原理。线性的、对称的关系适用于其结构透明的简单系统。在复杂系统中，必须找到打破对称并利用非线性放大能力的机制。这是由存在大量相互作用的层次和为资源而竞争来保证的。

社会系统是非线性的，也是不对称的。同样的一条信息对于不同的个体具有不同的效果，小的原因可能导致大的结果。社会系统的竞争本性常常受到权力关系的调节，从而保证了一个不对称的关系系统。有必要着力强调，这并不支持偏向于支配关系或利用关系。这只是一个承认复杂性的证据。非线性、不对称、权力和竞争是复杂系统的不可避免的组分。这正是驱动着它们的发动机。如果不是在婴儿与成人之间的不对称关系，婴儿便不会存活。如果不是在教师与学生之间的不对称关系，学生便永远都不会学到新的东西。

如果国家没有权力，它就没有理由存在。如果女人与男人完全一样，我们的世界便会毫无情趣。

这些考虑对于社会理论具有重要的含义。社会是通过不对称的权力关系而集合在一起的，这个事实并不意味着这些关系从来不被利用。恰恰相反，它们不断地被父母、教师、国家和男人在加以利用，同时也被子女、学生、公民和女人在加以利用。要点在于，对这些形式的利用的解答，并不落在其中能量均匀分布的某种对称空间之中。在非线性驱动的复杂系统中，不可能存在这样的空间。期望这样的空间能够以某种持久的方式被创造出来是不正确的[3]。为了与利用进行斗争，在此仅仅有一种选择：必须进入到网络的竞技场中。由于这种趋法的使用没有任何方式可以保证其成功，因此也就没有更高明的道德根基可言，无论我们是否拒斥这种现代主义伦理学的抽象规则。

（v）相互作用处在相当短的范围。复杂网络中的要素通常主要是与周围的要素相互作用。在大的网络中，这会导致群团或要素的聚集而形成某种特定的功能。利奥塔尔（1984：xxiv，61，66）把这种现象描述为“局域决定作用”。要素以其所能获得的局部信息来运行——它们必须如此，因为复杂（后现代）系统中没有对于信息流的元层次控制。因此系统的行为最好是以局域“话语”的多重性为特征。

关于重复的危险，有必要强调，这些局域决定的群团不是相互独立的。尽管直接的相互作用是短程的，但却绝不会预先排除大范围的影响。不同的簇群是相互关联着的，由于这些关联是非线性的，它们可能产生出大的效果，即使是相

互关联稀少时亦如此。重要的事件能够通过系统得到相当迅速的反响，但是它们绝不会以某种纯粹的形式传播，因为它们受到处在它们通过的簇群的不断调整。

（vi）**相互作用中有回路。**反馈是复杂系统的一个本质性方面。反馈不能以控制理论来简单地理解，而要作为大网络中的内在的相互连接的回路。这就意味着，要素的活动可以直接或间接地影响其本身。在后现代理论中，这将以自反性问题来表示（参见 Lawson 1985）。如果承认信息是通过系统而增加的，并处于不断的变换之中——由其他的信息或由其本身——那么就变得难以保证对任何信息片段的“真实”解释。信息只可能局域地加以解释，因此只有通过延异的动态性——反射于自身并转变着自身——加以解释。这些动态性排除了在元层次上定义真理或起源，因此被看作是后现代的困境——“一种我们的真理、我们的价值、我们最珍视的信念的危机。”（Lawson 1985：9）

这的确可能是知识的危机，但是，必须要强调的是，危机并非是诸如尼采（Nietzsche）、海德格尔和德里达这样的理论家的“颠覆性”分裂活动的结果。它是我们后现代社会复杂性的直接结果。当利奥塔尔坚持知识的状况是局域地决定的，也正是指出了这一点。自反性并不导致悖论，而只有假定所有的悖论都必须在元层次上加以解决时才是一个问题。如果必须保持在网络的层次上，就必须处理后现代状况的形似性。其含义不是不可能解释信息；而仅仅是意味着，所有的解释都具有偶发性和暂时性，都从属于一定的语境和一定的时间框架。

（vii）**复杂系统是开放系统。**我们已经指出，局域话语并非是封闭起来的，而是与他者处于相互作用之中。社会系统也与许多其他系统处于相互作用中，例如包括与生态系统的相互作用。这种关系已经受到了新的审视，导致强大政治团体将环境作为主要关心的问题。自然不再是人们利用的被动对象，而是使得人成为人的关系集的一部分。

（viii）**复杂系统在远离平衡条件下运行。**复杂系统需要不断的能量流，以作为复杂实体而变化、进化和生存。平衡、对称和彻底稳定则意味着死亡。正如能量流对于对抗熵并保持系统的复杂结构是必要的，社会也只能作为过程而生存。它不仅仅由其起源或目标来定义，而且也由其行为来定义。在后现代的社会中，这种不断的活动，这种平衡的缺乏——特别是通过大众媒体的作用——被推向了甚至更高的水平。这具有许多令人不安的效应，不容否认的是，人们必须发展起一定的技能以应对这些状况，但是，如果向往完全平衡的状态，那也就是向往石棺。

（ix）**复杂系统具有历史。**历史的重要性已经一再地得到了强调。有一点值得再次指出：复杂系统的历史不是某种客观地给定的状态；它是分布遍及系统的痕迹的集合，总是向着多种相互作用开放的。在这种意义上理解的历史肯定适用于后现代社会，这使得它几乎是无历史的。后现代主义的确拒斥对于历史的解释，这种解释将其拔高为目前状况的真正含义之门的万能钥匙。不过，对于现在的思考，如果不考虑到过去——或未来——就是不可能的。

（x）**个体要素并不知道它们所嵌入其中的整个系统的行为。**这

是一个更复杂的要点，需要仔细加以考虑。我们已经指出，系统中的要素仅仅能够对局域信息作出响应，记住这种信息可以是相当丰富的。我们也已经表明，复杂系统中的单个要素，不是由其本身而具有意义的，而是通过他们的相互作用的模式而获得意义。这里提出的这个要点，稍微有些不同。单个要素不可能包含着整个系统的复杂性，因此既不可能受控，也不可能被完全理解。

因为在后现代社会中可利用的信息是极其大量的，所以我们常常生活在我们所获得完整图像的幻觉之中。因为我们社会的复杂性，所以这是不可能的。由于我们通过我们的行动而部分地创造着社会，所以任何人都不会得到关于社会现在状态的完整图像。在这方面，所有的要素都同乘一条船。一些要素可能具有对于特定方面的更多控制——我们的政治模型仍然采取的是允许单个个体有太多的权力的方式。单个要素不应该，正常情况下也不会施行对于去中心化系统的完全控制。例如，我希望与通货膨胀进行战斗，但是我却无法测量自己的消费模式的效应。只有在我的行为与那些大量的其他经济行为者结合起来时，这些效应才是清楚的。

认为社会的结构是社会系统的一种涌现性质，可以创造出一种感受，即人们自己的活动是无关的或无意义的。实际上未必如此。首先，你的活动相关性是由你进入了网络竞技场中的有效性所决定的，不是由试图从上帝的观点来理解生命所决定的。其次，必须要记住，因为相互作用是非线性的，所以小的原因能够引起大的后果。不过，这也就意味着，我们行动的效果是有些不可预测的。

从这10个特点来看，清楚的是，从复杂系统的分布式模型来分析后现代社会是可能的。其中特别有趣的是此模型对于伦理学表现出来的含义。在本章的结尾处将继续讨论这个问题。在接下去的一节中，我希望考察此模型对于我们关于语言以及关于科学知识状况的理解的含义。

7.4 作为复杂系统的语言

在第2章中，我勾画了关于语言的两种范例性趋法：乔姆斯基的基于规则的趋法和索绪尔的关系趋法。我指出，联结论模型与索绪尔语言学之间是密切相关的，并讨论了复杂网络如何能够与德里达关于索绪尔的发展和批评(第3章中所讨论的)结合起来。在第5章中，我分析了表征问题，特别是参照了意义的表征理论。语言因此已经得到了很大的重视。在本节中，我希望小结复杂系统的分布式模型对于语言理论所具有的意义。为了真正以这些术语来描述语言，并以此为背景来对种种语言理论作出评价，就要求对其本身进行全方位的研究。随之建立起成熟的关于语言的联结论模型，并从计算上对它进行测试将是同样使人生畏的任务。我在这里仅仅希望表明，这样的方案为何是富有意义的。

类似于后现代社会所表明的获得复杂系统的重要特点的方式，我们能够通过对语言的考察以确定在何种程度上此模型也是适用的。因为我们在先前已经谈论了这个问题，所以这里将较快地向前推进。任何适当的语言，都是由大量的词汇组成的，词汇的意义是通过它们相互之间的关系而构成的(当然，这是在索绪尔的范式之中)。在这些关系中，非线性和不对称至关重要。例如，尽管“红色”的意义是通过相对于系统中所有其他词汇的一套关系而构成，但是它相对于

"蓝色"和"黄色"的关系，并不相似于它相对于"血"或"物质论"的关系。语言关系的对称，也是比喻成立的先决条件。比如说，A 是 B（他是猪）并不意味着 B 是 A。不过，大多数术语都主要是由它们相对于直接相邻关系（"红色"是由其相对于"黄色"、"蓝色"、"绿色"和"紫色"等的关系）而组成。长程相互作用是以中介为媒体的，常常导致比喻。当比喻被固定下来，就意味着网络已经转变其结构到了词与词已经变得互相邻近（在网络拓扑的意义上）这样的程度。

语言的自反性——网络中的回路——可以由延异的动力学来作为例子。语言，以其被运用的方式而不断地转化。例如，一个术语被运用的频率，能够引起该术语或是限制其语义领域（因为我们已经过分地习惯于它），或是扩展其语义领域（因为它被不适当地运用或被运用于不同的语境中）。该术语的运用从而变更了术语本身的意义。由于语言是一个开放的系统，它以多种方式与环境相互作用。环境的新的方面将对语言具有效应。已经长期没有异议的词汇（例如"chairman"），可能突然变得有争议。我们还被迫地，尤其是在后现代世界，不断遇上关于新事物的新词汇。

如果语言被封闭起来，如果它被形式化为一个其中意义固定不变的稳定系统，那么，它就将死亡，或者开始走向死亡。活生生的语言是处于远离平衡态的。它在变化之中，它处于与其他语言的接触之中，它在被误用并发生着转化。这并不意味着意义是一个随机的或人为的过程。这意味着，意义是局域现象；在一定的时间和空间框架中才是有效的。因为所有的语言也都具有历史——对于术语的意义共同负责的历史——意义比人们所想像的要更为稳定，甚至在价值流动并繁殖的模型的语境之中也是如此。词汇承载着不能想当然

地被放弃先前意义的痕迹。首先，语言是一种系统，其中个别的词汇并不具有其本身的意义。只有当个体词汇在系统游戏中被捕获了的时候，它才产生出意义。

意义的表征理论的问题，在第5章中已经详细讨论了，这里不再重复讨论。那里的最后结果，是拒斥了其中将句法层和语义层割裂的理论。这不仅是由于其后结构的视角，而且也结合了普特南对经典表征的批评。不过，我的确希望返回对于作为自组织系统的语言的讨论。

那些支持对于语言的后结构理解的人，常常面临着如下的问题：如果语言是由一组能指之间的关系构成的，那么如何来解释在语言和世界之间的关系呢？这里关心的是，如果意义被归因于仅仅是语符的运用，语言将变得漂浮不定，以致再也不可能有关于经验真理的断言。尽管这种反对意见通常是由怀旧情绪导致的，念念不忘能够调节意义产生的连贯的元叙事，但是它的确提出了一个值得仔细考虑回答的问题。

我已经指出，词汇和对象以直接方式一一配对——经典的拟态表征——是不可接受的。它没有充分承认这样的事实：语言是一个复杂系统。它假定了存在着客观的、外部的观点，并祈求于实施这种“配对”的行为者的特性。语言与世界之间的关系，既非直接的透明的，也非客观上受控的，但却是这样的关系——没有它，自然语言就不可能存在。把语言理解为自组织系统，我们便能着手勾画出关于这种关系的更精致的理论。

自组织描述了一个复杂系统何以能发展和改变其内部结构。此过程受到为系统资源而竞争的行为驱动。从环境来的信息进入到系统中(通过某种感受机制)，并与已被编码并贮存于系统中的信息发生

相互作用，引起系统去适应和改变它对于环境的响应。所有的这些都已在第6章中进行了描述。我们在那里聚焦于将大脑和神经网络作为自组织系统，但是这对于语言也同样是成立的。

我们这里谈论的要点在于以下这个事实：从环境来的信息对于系统具有直接的、尽管是非决定论的影响——它引起了系统中的一定的变化，但它并不完全地决定变化的本性。从环境来的信息是以非线性的方式与贮存于系统中的信息发生相互作用的(记住，系统的记忆是分布式的，而不是图标式的)。进来的信号受到系统的历史的调节，其方式是，它结合了重要的新方面，但抵制不必要的涨落。在任何给定时刻的系统状态，是其环境、系统的历史以及系统为发挥功能必定对其环境产生的效应的条件的共同结果。

事实上语言是如何与其环境(“外面的世界”)相互作用的？主要是通过语言的使用者，他们为了在环境中生存和运行而必须与其相互作用。事实上，语言是我们用于此任务的最重要的工具之一。我们试图使我们的经验具有意义，并在此过程中创造和扩展我们的语言。这不是抽象的心智作用的结果，也不仅仅是对环境反映的结果。有意义的语言在时间推移中通过自组织而进化，滞置在主动和被动之间，其中，语言的有用的或有效的形式存活了下来，而过时的形式便腐朽了。一个例子可能有助于我们说明这一点。“cow”这个词(或某个等价的词)，可以通过不同的机制而进入到某人的语言中——包括下定义(通过运用已知术语来表明其关系)和形象展示。不过，这些机制都不足以由其本身作为某个适当语言理论的基础。“cow”这个词将获得的意义，其重要程度与牛对于语言的使用者的重要程度相当，以及，当然还有与已经成为使用者语言系统的部分的那些要素发生相互作用。此术语被与其他已有术语进行比较，以满

足当前的约束。此术语最初也许是相当不精确的(“相当大的、黑白相间的动物”)。术语的意义将随着该术语的重要程度而得到发展。例如，如果只是为了在一个重要国际展示会上作为种牛进行评判，术语“cow”将被分化成大量的细节。如果是一个城市的小孩，此术语仅仅有一些微弱的内涵诸如“牛奶”、“牛排”以及一些并不真正为人所理解的侮辱性用法。

如果语言能够被描述成自组织系统，在语言和世界之间的关系问题，就以相当精致的方式得到了解决。世界对于词汇的意义具有直接因的影响，但并不决定词汇的精确意义。意义从复杂的相互作用过程奔涌而出，这种相互作用的两方面，一方是源自世界的信息，另一方是由先前的相互作用所致的已存的关系之网。这使得语言成为一种活生生的、进化的系统，能够应对大的复杂性。如果环境的某些方面具有大的重要性，系统就将组织自身向鲁棒而精确地解释这些方面的方向发展。在资源分配上，系统不会过多地分配给那些不被频繁使用或没有太多利益的术语从而造成资源的浪费。对语言的这种解释，允许我们去为痕迹和延异的动力学发现重要的场所；也会导致我们承认，语言系统将组织其自身到某种临界点，最大量的意义在此将产生出来，而不致游离于世界而漂浮不定。

语言和世界之间的关系总是有趣的，当我们从事科学时就更加有趣。那么，我们就负有责任，在条件允许下使得这种关系尽可能地清晰。因此，让我们从复杂系统的视角，考察一下科学与后现代理论之间的关系吧。

7.5 从后现代视角理解科学知识

如果我们无需借助某种元位置，而是根据帮助我们应对世界的复

杂而鲁棒的系统来描述语言，那么，我们就应该能够对于科学知识的境况进行一些类似的探讨。我将再一次地运用利奥塔尔关于后现代社会中科学知识状况的分析，但是从引用维尔顿(Wilden 1987：309)来开始这个讨论。

> 对于简单性的反抗使得被忽视了的复杂性主题，或更精确地说，有组织的复杂性(受约束的多样性)，成为值得科学家——生物学家、生态学家、哲学家、人文主义者、文化批评家和其他人——注意的主题。
>
> 一旦我们都认识到我们可能结束核攻击或令人作呕的东西，倾向于多样性的多层次的革命就正在到来。
>
> 我们长远的未来——如果我们有的话——现在完全取决于掌权的人们开始理解新科学的基本事实：摧毁着环境的开放系统最终会摧毁其自身。
>
> 有组织的复杂性，是行星地球上的生命、自由和新颖之源。

首先，维尔顿提供了一种复杂性定义，这个定义与在本研究中发展起来的关于复杂性的理解是一致的。他强调了组织作用原理，以保证复杂性不会与混沌相混淆。复杂系统是受到约束的，它们具备有组织的结构，但是在那些约束下，系统必须最大限度地多样化。曾经被忽视的关于复杂性的研究，现在能够以科学的方式进行探索。

维尔顿提出的第二个要点，涉及这种科学趋法的本性。他不是把科学看作某种处于封闭的、破碎社会系统中的一个孤立要素，而是强调了对于复杂性科学的政治重要性和社会重要性。

我们需要把握复杂性，以保证我们的生存。同时，复杂性是自

由之源。因此，在科学和政治学中，承认复杂性都是至关重要的一步。自由和正义并不能以某种中心控制的形式来强加上普遍规律而达到，如果科学保持为一个封闭店铺并言说着自己的私人语言，也就不可能有科学的繁荣。科学活动需要更加开放，不仅指向世界，而且需要指向内部。种种科学学科之间的障碍需要被跨越。

利奥塔尔明确地承认了这些要点，他还指出，科学知识传统上被内部地合法化，即其标准是，若某事物有资格被称作科学的，则这是由科学本身来决定的。利奥塔尔(1984：25～26)评注了这种保守性趋法的5个性质，他将这种趋法称为“科学语用学”：

(i) 科学知识需要一种语言策略——指示对象(denotation)被强调而其他的都被排除掉。尽管其他类型的陈述，如在此过程中突然出现质疑性陈述或说明性陈述，那它们也仅仅是提呈了“辩证论证中的转折点，这必定终止于指示性陈述”(25页)。一个科学家，就是能够作出这种陈述的人——这种陈述是可证实或可证伪的，只有此领域中的专家才能理解。

(ii) 以这种方式产生的科学知识不是一般“社会契约”的一部分，而是将自己组织进排外的机构和体制中的专家和职业人员的财产。“知识和社会之间的关系……成为一种相互的外在性。”(25页)

(iii) 研究过程中起作用的只是研究人员的能力。(人文科学中)能力无需从消息的接受者或研究的主题那里获得。“一个人无需去了解何种知识告知他应当如何。”(26页)

(iv) 科学陈述并不由于被报道或变得流行而增加其有效性。任何此种陈述的有效性，都必须得到证明。

（v）科学知识是积累的。科学家被假定掌握着所在领域的知识主体，并仅仅当获得了不同于先前的新陈述时才加入其中。

这种知识——科学知识——与更一般的知识形成了鲜明的对照，利奥塔尔将这更一般的知识称作“叙事知识”。叙事知识的语用学也是以一些性质来加以描述(18～23页)：

（i）叙事知识的标准是柔性的，并在叙事起作用的社会中被动态地定义。

（ii）在这些叙事中没有特定的语言形式具有优先性。它们赋予自己以语言游戏的多样性。

（iii）在这些叙事的“传递”中，发送者、接受者和所讨论的主题都以加强社会契约的方式来考虑。

（iv）叙事具有奇怪的时间本性。它们主要不是回忆过去，而是使得过去的事件与现在的事件重新发生作用。叙事的意义不在于它得到某种重要历史片段的支持，而是在于它的现在诉说的节拍和韵律。

（v）不需要任何特定的程序来“权威化”叙事过程。叙事者不是孤立的，而是在发挥着整合者的功能，所有参与者都能够在可能的角色中(叙事者、听叙者、英雄，等等)找到他们自己。

利奥塔尔所提出的这两种形式的知识——科学的和叙事的——已经如此长期地分离，以至于它们已经变得不可通约了。叙事知识可以包含着科学知识的某些部分，可是科学知识却是独立地合法化。叙事陈述不可能服从于论证和证明(27页)。然而，这两种知识的分

离，导致了元叙事被看作是可疑的后现代时代中科学知识的合法性危机。不过，科学元叙事的衰落并不仅仅是某种理论趋法的结果，而是科学现在必须面对的多样性和复杂性的必然结果。

> 科学知识"危机"的迹象自从19世纪末起便开始积累，这不是科学知识偶然增殖的结果，而是技术进步和资本主义扩张的结果。或更准确地说，它代表着知识合法性原理的内部腐蚀。腐蚀在猜谜游戏的内部起作用，并通过松开其中每一种科学都有其自身位置的百科全书网的编织，最终使得它们获得了自由。(39页)

这种新的自由不仅对于科学与社会之间的关系是重要的，而且对于发展科学知识的新理解以及科学本身的实践也是重要的。在科学学科之间的传统的划界，不再具有不受挑战的地位了。事实上，当代科学的最有趣之处即是学科的交叉：生物技术，遗传计算算法，计算机网络和虚拟现实。"学识的思辨等级性，让位给了固有的——过去是这样——探索领域的'扁平'网络，相应的前沿也处于不断地流动之中。"(39页)[4]

利奥塔尔这里赋予科学探索两种重要的特点。首先，它是固有的，即条件性的(contingent)。这意味着，如同在叙事知识的情形，其价值是由游戏的参与者按照其现实的需求和约束所决定的。其次，探索的领域关联在"扁平"网络上。我们已经意识到把后现代社会理解成网络的重要性，但是我们必须记住这种网络是"扁平"网络。它不受某种下面的基础的支持，也不会通过来自上面的某个一般抽象而保持在一起。我们只可能通过复杂的、扁平漂浮的网络而追溯种种的叙事途径。

我们现在拥有了去发展对于科学知识的“叙事”解释的框架。科学知识，不是指示性的、外部的、自我中心的、逻辑的和历史积累的，而是通过相互作用和多样性而产生的，并越来越嵌入到更大的社会网络之中。科学，也只有加入这个网络竞技场中才可能生存。

有用知识的标准，不再是指示性，而更多的是灵活性。因为它形成了开放系统的一部分，所以必须从更广泛的视野中来考虑问题。它不可能仅仅依赖于权威来使自己合法化，无论是历史的权威[5]，还是专家的权威。科学知识，如同社会网络，也是以这样的方式将自己组织起来的：保证那些产生差异的叙事——以条件性发生而非抽象的术语——永久地进行。不用像某些现代主义者那样感到害怕，我们没有理由担心这比遵循某种保守趋法将导致更少或更不可靠的知识。

对于叙事知识也是科学知识的思想，现在可以作一小结。我们所生活的世界是复杂的。这种复杂性是多样的，但却是有组织的，而非混沌的。对它的描述不可能还原成简单的、连贯的并普遍有效的话语。如果我们从网络来对复杂性建模，任何一定的叙事都将形成一种通过网络的途径或轨迹。这类途径具有多种多样性。网络不仅是复杂的，而且是动力学的。正如我们通过它来追溯种种叙事途径一样，它在变化之中。不过，所有的途径都受到网络的局域结构的约束。这些约束在某些地方也许是相当松散的，而在另一地方它们也许是相当紧凑的。有许多叙事途径可以追随，甚至处于两个特定的点之间，但这个事实并不意味着怎样都行。所有的叙事都从属于某种形式的约束，一些途径会被剔除掉。所有的途径都共享着偶发性(contingency)和暂时性的特点。从策略上考虑，网络的一定部分可以关闭起来并固定下来。这种“架构”的过程是科学探索的必

要部分，但是由此产生的知识却保持着与该框架的联系，不可能在时间或空间的意义上被一般化。

在这一点上，有一些其他的思想家，他们对于科学的理解类似于来自利奥塔尔的理解，在此将扼要地提及他们。布莱克韦尔(Blackwell 1976)建议了一种框架，他称作“对科学理论的结构主义说明”的框架。他没有提到任何的后现代或后现代结构主义作者，但是承认他受惠于皮亚杰(Piaget)(264 页)。然而，他的论据却拥有为他们所熟悉的声调。他写道：“西方哲学中几乎所有的知识理论都是以存在认识论终点的思想为基础的。”(263 页)作为例子，他援引了柏拉图(Plato)的形式说，亚里士多德(Aristotle)的四因说，笛卡儿(Descartes)的天赋说，洛克(Locke)的初性说，休谟(Hume)的单纯印象说，康德(Kant)的感性和知性说，皮尔斯(Peirce)的科学理想状态说，胡塞尔(Husserl)的本质说，罗素(Russell)的感觉材料说以及卡纳普(Carnap)的世界逻辑原子功能块说(263～264 页)。布莱克韦尔然后否认存在认识论的终点，提出了更为温和的观点，承认“我们个人的经验以及可利用的积累证据都是局限于相对狭窄的空间和时间范围的”(264 页)。在说明了我们的知识以后，我们总是必须从发现我们自己的地方开始的：处在川流不息之中。为了应付这种事务状态，他提出了“结构主义者”说明。

追随皮亚杰，他将结构定义为“任何有组织的受规律支配并是自调节的转换系统”(266 页)。他强调，结构是“转换系统而不是部分或要素的系统”(267 页)。系统的自我保持(可解读作自组织)目标在于保持动力学结构，而不是保持任何特定要素的存在(268 页)。在发展科学理论中，人们不应该落入鉴明基本要素并进而询问它们是如何结合形成理论的诱惑之中。“这正是错误的趋法，而在我们看

来，这是当代科学哲学在试图提出对于理论的适当说明时遭致失败的主要缘由。”（269 页）

按照他的“中流原理”（mid-stream principle），理论没有终极要素，仅仅具有中介要素。人们不应该集中于关注要素，而应该集中于关注系统。理论因此“被设计来作为过程，而不是静态的、逻辑的形式表述”（269 页）。尽管他并没有涉及复杂性或对于多样性的需要，但是布莱克韦尔表明了对于科学理论的条件性以及自组织的作用的敏感，这里的自组织是一种在适应新环境中允许系统结构具有连续性的过程。

赫西是另一位对于传统的分析科学哲学有着理解的思想家（Hesse 1992：49）。她的论据是在英美哲学的语境中提出的，尽管其中并不包括结构主义或后现代思想的要素[6]，但她对语言及其提供的接触世界的方式的复杂本性一直有兴趣。她指出，比喻是“语言的基本形式，并先于（历史的或逻辑的）文字”（54 页）。对于比喻和类比推理的强调，将科学哲学所关注的重点从归纳推理转向了问题的范畴和簇群的划分（Hesse 1988：324）。她运用维特根斯坦（Wittgenstein）的族相似性指出，我们获得的认知，不是源于对特殊情形下作为支柱的所有逻辑关系的无穷尽计算，而是藉发现了充分多的类似，并将这种具体的关系运用于已经满意的其他关系。

将具有相关特点的事物聚集在一起，当然也正是自组织神经网络的作用。尽管赫西并没有明确运用复杂性的联结论模型，但她的确在使用“类似的网络”（324 页）概念做工作，而且她提出的许多要点，都是与复杂性的联结论模型相通的。对此更为清楚地体现在，她赞成一种温和的科学实在论，赞成知识“是特殊的而不是一般的，局域的而不是普遍的，近似的而不是精确的，立即可描述的和可确证

的而不是理论深度的和还原的”（Hesse 1992：53）。

劳斯（Joseph Rouse）在他发展的一种具有叙事结构的科学哲学中，明确地采纳了后现代要素。他的中心主张是提出：“科学知识的可理解性、意义和辩护，源于它们已经从属的、由正在进行的科学研究的社会实践所提供的、处于连续重构的叙事语境”（Rouse 1990：181）。不幸的是，他并没有集中精力发展这种“科学的叙事重构”。相反，他特别参照现代话语持续于科学哲学中的方式，试图勾画出科学哲学的领地，以领会谁适合于何处（Rouse 1991a）。在这方面，劳斯作出了如下的评注：

> 因此，我使得相当新近的科学哲学和科学社会学的现代主义精粹特征转向了处在全然可以两者择一之中：要么是在实在论或理性方法论的一方，要么是在以相对主义和戳穿所谓的科学文化霸权的另一方。（Rouse 1991b：608）

从这里出发，他得出结论，激进的后实证主义科学哲学家，如库恩、拉卡托斯和费耶阿本德实际上仍然属于现代性。他承认，这初听起来可能会觉得奇怪，但是按照哈金（Hacking）的说法，如果我们不把“现代性”看作一种地位，而是看作一种“共享的冲突领域，在此必定有使得鲜明的、相应的不赞同成为可能的大量基垫证据”（610页）。后现代的科学哲学必须脱离这种“要么／或是”的结构，但是劳斯并没有真正阐明如何来做到这一点。

劳斯（1990）在一篇稍早些的文章中——尽管文中并没有直接参考任何后现代思想家——承认了受惠于海德格尔，的确在其中清楚地表明了他对于科学知识的一般性的理解。他在文章中关心的是有关科

学叙事的连贯性的一个棘手问题。一方面，劳斯拒斥了科学统一的思想，“它假定了一种认识的理想，即把关于世界的科学表征统一为单一的包揽无遗的连贯的图景”（192 页）。他通过否认科学的目标是“对于研究活动抽象所获的表征体之积累”（192 页），而作出了这种拒斥。科学的进展是离心的而非向心的，在此过程中，已建立起来的结果往往被置于一旁。科学关注的，是实践的而不是表征的，所以尽管有对于连贯性的关注，但这仅仅是局部的，甚至比可应用的学科层次还要局部，更不要说与科学的整体相比了。

另一方面，劳斯认为存在一种有差异的、重要的观念，其中科学知识是不统一的。这可由这样的事实所表明，在学科之间的边界常常被打破，不仅是由于科学家为了工具和概念而掠夺其他学科，而且也由于杂交学科(hybrid disciplines)的形成，杂交学科随着不断增加着复杂性的问题进入研究视野而在不断发展。

> 我们所获得的，不是单一的连贯的世界图像，而是一个结合起各种科学努力的相互关联的日益复杂的网络。……连贯性的丢失，常常愉快地从新探索的可能性、以及从事这种探索的新能力的创造中得到补偿。（192～193 页）

劳斯对于竞争的敏感性，也是值得一提的：“一种‘技术’，只不过是对于冲突历史的一种反应”（188 页）。设定了这种框架，我们可以得出结论，尽管劳斯并没有就后现代社会中科学知识的实际动力学进行阐述，但他不会反对我在解读利奥塔尔中提议的“网络竞技场”。

不过，在他最新的著作中，劳斯(1996)几乎没有再对后现代科学说上一点什么。除了有一点关于性别问题的文献之外，没有对科学

实践中的伦理学维度进行探讨，也没有涉及大陆的理论。像利奥塔尔和德里达这样的思想家只是顺道提及。而对于我来说，阅读这种书就如同将其咀嚼一般地愉快。

我希望讨论的下一位理论家也显得令人失望，至少就本研究中多采取的趋法的语境而言是这样。保罗·丘奇兰德是关于神经计算趋法具有重要性的强烈支持者。他和帕特里夏·丘奇兰德属于最早探讨神经网络哲学含义的哲学家之列(P. M. Churchland 1984, 1989; P. M. Churchland and P. S. Churchland 1990; P. S. Churchland 1986; P. S. Churchland and Sejnowski 1992)。主要问题不在于他们是否接受或利用联结论模型，而在于他们是如何进行联结论建模的。我们可以考察《一种关于神经计算的观点》(A Neurocomputational Perspective)，其中保罗·丘奇兰德(1989)考察了神经网络对于科学知识理论的重要性。

在一篇冠名以"可观测量的本体论地位：超经验的优点"(The Ontological Status of Observables: In Praise of Superempirical Virtues)(139～151页)的论文中，他将自己定位于实在论者，断言"对于所有层次上的认知，甚至对于可观测层次上，理论的全局性最优是真理和本体论的最终尺度"(139页)。他的实在论，比起我从这段话中(141页)推论的要更加慎重，但是仍然承诺了这样的思想，即有一个独立于我们的"认知"而存在的世界，我们建构起关于这个世界的表征(151页)。因为不同的表征是可能的，它们必须被比较，并从中选出最好的。选择不可能以"经验事实"为基础来实现，而"必定是以超经验基础，诸如相对连贯性、简单性和解释的统一性来实现"(146页)。这就应该清楚了，从这种见解来看，他没有打算探讨偶发性、复杂性和多样性。

丘奇兰德的趋法并非是朴素的或不精致的，这可从其核心论文“论理论的本性”（On the Nature of Theories）（153～196页）中看出来。按照他的观点，关于理论的经典观点认为，理论是“一组可在一阶谓词计算中表达的语句或命题”（153页）。在那种观点中，理论是可陈述的。只有这种理论可以执行理论的基本功能（即预见、解释和在理论间的归约）；它们允诺了（以形式术语来）说明学识本性和理性，其最终的价值是真理（153页）。丘奇兰德拒斥了把句子和命题的态度看作是“认知物所使用的最重要的表征形式”（158页）的观点，并提出人们倒是应该对大脑的“表征和计算”（159页）进行考察。在一个关于神经网络（或平行式分布处理）的引介之后，他讨论了后者对于理论本性的含义。他的确指出的一个重要方面是：神经网络在做什么的解释只可能在权重层面给出（177页）。对更高层次（逻辑）解释的需要因此被消除了[7]。对网络行为进行解释所必要的一切，都包含在权重的（材质）价值之中。

尽管这一点的确与本研究中的见解是相同的，但是丘奇兰德却完全是以静态的方式来理解的。一个解释是由特定时刻集体的权重值来提供，即“权重空间的一个点”（177页）。他的这种表征倾向也许并不令人惊讶，但是他的确低估了把解释和理解作为过程的诠释。这也许部分是由于丘奇兰德如同我一样，使用了简单的、单向的“后向传播”网络来解释神经网络的工作情况。尽管他承认归复式网络的重要性（208页），他却没有以这些网络的动力学为基础进行扩展。因此，我没有发现他有对于自组织过程的分析——这是一种以归复为前提条件的过程。

他从网络模型中提取的另一个重要的结论是，所有的经验观察都是负荷着理论的。因为信息进入此网络只可能从已有的权重模式来

进行处理，“任何认知活动都不是在没有某种理论或它种理论的情况下发生的”（188 页）。这种观点接近于本研究中所采取的观点，但是他仍然将理论形成的目标看作是对于权重空间加入“有用和良好结构的子部分”（189 页）的发展[8]。对于这些分割，没有任何先验的东西，因为丘奇兰德意在放弃真理概念，“在追求某种甚至比真理更有价值的认识目标”（150 页）。不过，他并没有提到这种目标究竟是什么。

总之，丘奇兰德运用神经网络来作为一种解释工具是重要的一步，但是在我看来，进一步则被他的连续地承诺表征之重要性而严重地阻碍了。他并不相信存在着单一的、正确的表征，并相应地对于会聚实在论(convergent realism)是不抱希望的(194 页)，这必定使得他非常想放弃表征概念。不过，他坚持这样的思想，即世界必须“被割裂”（carved up)，网络通过发展起表征而进行这种割裂。他的最后的关于信念的文章仍然如此：

> 坚持由我们的最好理论所提供的表征，对于我们最好和最敏锐地把握现实仍然是成立的。理论的全局最优，仍然是理性本体论的基本尺度。而这一直是科学实在论的中心主张。(151 页）

最后，我希望提到一本书，它构成了一个对我并非是截然不同但结论却相反的方案。阿吉罗斯(Alexander J. Argyros 1991)在《一场幸运的风暴：解构、进化和混沌》（A Blessed Rage for Order：Deconstruction，Evolution，and Chaos)中，考察了德里达的工作，然后将其抛弃。他的趋法是一种保守的趋法，按他自己的看法是“一种基础主义”（2 页）。他坚定地持有进步的理想(323－331 页)，偶尔也热心而虔诚地说：“如果我们以某种更好的东西取代皇

帝[没有穿衣服的皇帝——是一个现代保守主义分子所热衷的比喻]，那么我们必须相信，并且是深深地相信，更好的选择是可能的。”(5页)以这种趋法为前提，他集中关注于混沌理论，将其作为引导我们实现希望之旅的科学，也就不意外了：“[混沌理论]能够作为一个开端，从而结束400年之久在科学和人文之间的分裂。而且，混沌还允许我们重新肯定这样一些遭受攻击的概念，如普遍性、一致性、意义、真理和美。”(7页)

阿吉罗斯书中的基本前提并非真的与我的论据不一致。他主张，由自然科学所揭示的自然世界，当然“引导着人们对于文化和自然的解释，正如人们对于文化和自然的解释受到社会历史的压力所引导一样”(2页)。尽管这种主张看起来有某种循环的成分——人的解释受到由科学揭示的世界的引导，而科学其本身当然包含着解释——我的争论并非针对这种主张，而是针对阿吉罗斯关于德里达将会对这种见解表示异议的结论。

他对于后现代结构主义的理解，显然受到美国文学理论格调的影响，此种理论乐于将后现代结构主义看作某种无政府的、摧毁性的东西。正是这种严重有局限的对于德里达的理解，导致了他主张“解构，以及所有的反基础主义，都必须托起自然世界的全体性，包括它组成人类非精神性之类(non-mindlike)的组成部分”(2页)，以及“对于解构主义者，任何与自然的对话都只能是伪装了的独角戏”(3页)。因此，这里清楚表明了，对于阿吉罗斯来说，后结构主义是一种唯心论，它否认了任何关于人体以及关于世界如其所是的重要性。我所不能同意的，正是这种对于后结构理论的理解；我还设想，这也是德里达所不能同意的。很遗憾的是，阿吉罗斯选择的用来进行详细分析的文章，一般认为是德里达的一篇比较逊色的文章——关于评

论种族隔离的论文[9]。

让我复述一下第1章中陈述的观点：当我们处理复杂的动力学系统时，并不一定要赞成后结构的见解。我发现，这是极有帮助的，因此是重要的，但也不可能从其他视角得出与此类似的结论。不过我感到，在阿吉罗斯那里，他的基本保守主义，迫使他将潜在的盟友转变成了对手。

在本节中，我已经从联结论模型的语境中探讨了科学知识的本性。我的中心论点是，我们对于世界的描述需要具有对于复杂性的内在敏感性。通过对于一些理论框架的考察，已经清楚的是，后现代和后结构理论的确具有这样的敏感性，而且进一步说，从联结论角度来解读这些理论，提示了模式和约束概念都是内在于后结构主义的。条件性约束的存在，限制了对把这些理论解释成意味着"怎样都行"的反驳。我还已经表明，接受联结论模型并不由其本身导致对于复杂性的更深刻理解。这就是为什么我坚持认为，把复杂性理论与后结构主义结合起来将会是富有成果的原因。

7.6 后现代伦理学的复杂性

我希望再返回到利奥塔尔(1984)的主张，他对于后现代状况的分析，为我们提供了"一种策略的纲要，对渴望正义与渴望未知两者都将予以尊重"(67页)。利奥塔尔浑然不顾那些感觉到全然缺乏元描述(或规定)使得后现代状况完全是非伦理的(unethical)人的反对。他们的反对意见是，不由客观标准所支撑(或至少是由某种形式的一致意见所支撑)的任何形式的批评，都可以被轻而易举地解除掉。利奥塔尔挑选哈贝马斯作为采取这种趋法的人物，但认为这样的趋法是"不可能的，更谈不上是谨慎的了"(45页)。

利奥塔尔主张，哈贝马斯的趋法是由他所称的“论证商谈”(dialogue of argumentation)所构成，这里取决于两个假设：第一个假设是“所有的言说者都就规则或元规定对于所有语言游戏都普遍有效达成一致是可能的”；第二个假设假定了“对话目标都是共识的”(65页)。利奥塔尔发现，这两个假设都是不可接受的，主要是因为它们否认了后现代社会的复杂性；利奥塔尔对其本性的描述如下：

> 这是一个由种种变形意见种类(指示的、说明的、述行的、技术的、评价的，等等)的网络交织而成的怪物。没有理由可以认为，它有可能决定所有这些语言游戏都共有的元规定，也没有理由认为可修订的共识(如科学共同体中一定时间的有效共识)可以包括对循环于社会全体之中的陈述总体进行着调节的元规定的总体。结果，合法的叙事——无论是传统的还是“现代的”(人性的解放，思想的实现)——将当代的衰落与这种信念的被放弃联系在了一起。(65页)

哈贝马斯的趋法的第一个假设是直接与利奥塔尔强调异质话语的增加和形似性的作用相左的，而第二个假设则与利奥塔尔坚持认为意见纷争的重要性相左。共识并不总是不可能的；它是可能达到的，但仅仅是一种局限于时间和空间的局域现象。共识作为一个目标，将试图把社会系统冻结成某种特殊的状态。既然这是不容易达到的(也是不希望的)，那么更加好的(以及更正确的)策略是，发展起对于这种社会转变过程的敏感性。这也许表明，“共识的价值已经成为过时的和可疑的”，但是利奥塔尔主张，“公正作为一种价值既不过时也不可疑”(66页)。

从后现代社会的复杂性来看，公正的概念必定是一个有问题的概念，但是利奥塔尔承认有两个重要性，如果可预测，那么策略是：承认语言游戏的异质本性；并承认所有关于任何话语规则以及在该话语中被允许的“策略”的共识，都必定是局域的，换言之，“由当下游戏者一致同意，并注定最终会消失”（66 页）。

这种主张，勾画了一种能够以如下方式得到最好理解的实践的公正理论。认识到所涉语言游戏的规则成为所有游戏人在任何散漫的实践中都负有的责任。这些规则是局域的，即“局限于时间和空间之中”（66 页）。遵循这样的规则，人们必须承担起对于规则自身以及对于特定实践的效应两方面的责任。这种责任，不可能转变成任何普遍性指导原则或制度——无论它们是国家、教堂还是俱乐部。

在这里必须提出如下的问题：行为能完全与被称为“伦理的”一组抽象而普遍的规则协调起来吗？这里的要害是“伦理”一词的确切意义。作为现代主义梦想的一部分，它试图建立起一组能够在所有情形下调节我们行为的普遍规则。就其本身而言，这是一种高尚的理想，但是如果我们希望人类是由其伦理行为所构成的，那么我们就遇上了问题。只要遵循一组普遍规则（假定这样的规则是存在的）就不必涉及决策或两难，而仅需要计算。在一定情形下，规则将如何规定我的行为？这能被叫做“伦理的”吗？何种人类将这样地行动呢？显然，某些种类的自动机——其本身是由合理的——以规则为基础的原理所构成。现代主义伦理学，以及按照基于规则的系统对于语言和精神的理解，倒是都能够完美地适合它。

这与鲍曼（1992，1993）对于现代和后现代伦理学的分析相吻合。在他看来，现代主义建构起我们的存在的企图导致了不亚于将我们监禁起来的后果。而后现代的态度则使我们获得自由，不是按

照我们的喜好去做，而是伦理地行动。他承认，这里涉及一个悖论："在把全部的道德选择和责任归还给行为人。剥夺了行为者享受曾经为现代自信所承诺的普遍指导性原则的权力和舒适的同时，……道德责任伴随着道德选择的唯一性而生。"(Bauman 1992：xxii)实际上，这正是对于成为本章中心思想的此原理的另一种表述：人们不可能逃脱网络的竞技场。

在实践中，这种原理是如何影响道德行为的，它真正意味着所有的伦理原理都是如此偶发、如此短暂的吗？这种难题，在D·康奈尔(Druscilla Cornell 1992)的著作中有所探讨。她的目的是要"建立起这种有局限的[她使用这个术语来描述自己对于后结构理论的解释]哲学与伦理学、公正和法律解释的关系"(1页)。她集中关注的是阿多诺(Adorno)、德里达和列维纳斯的研究；尽管她主要是涉及法理学领域，但我的意见是，本书中提呈的大多数论据都将支持她的见解。

康奈尔肯定支持我们都是由一组复杂关系构成的这个概念，她以黑格尔(Hegel)作为重要的参考点，分析了对于这种见解的不同解释。黑格尔在对康德的抽象观念论的批判中，认识到我们是在社会系统之中构成的，但他认为这是一个最终将被完美地组织起来的系统。所有的人都将得到他们的特定位置。尽管黑格尔认识到，组织的(辩证)过程仍然在进行之中，但他在最后仍将其看作是一个保守的过程。到某个阶段便不需要进一步的转变了。阿多诺指出，人与人之间总是存在着无法被还原成加和性系统的差异。他由此提醒我们差异的重要性，这不是某种阻止我们在系统中发现一个舒服位置的东西，而正是构成了我们人类的东西。

作为一位分析了系统理论的伦理学含义的当代思想家，康奈尔(116～154页)讨论了卢曼(Niklas Luhmann)的著作。卢曼将社会看

作是复杂的自组织系统。因为我们是这个系统中的一部分——我们实际上绝不会站在这个系统之外——我们没有别的选择，只能承认这个系统将能以最好的生存方式来组织自身。此系统将“进化”，但不可能被“转化”。我们除了呆在系统“之中”外别无选择。例如，现在对于法律原理的理解因此是正确的理解——一个保守的、甚至是实证主义的结论。这是一种“弱”的后现代主义，是阿吉罗斯反对的一种建构主义(参见前面)的例子。这也许使他感到吃惊，康奈尔正是利用德里达的观点来反对卢曼把社会理解成自我约束的、完满的系统。

康奈尔和德里达论据的关键在于，他们关于系统中时间流的解释(128 页)。关于系统的时间本性的传统解释，包括卢曼的见解，给予了现在优先权。延异概念的极好教益在于，它提醒我们，当我们试图建立起(例如)事件的意义，不仅要考虑到过去，而且因为我们不可能完全地预见事件的未来效果，所以还必须考虑到未来，而不论我们对于未来究竟如何地全然不知。于是，德里达不是猛地举起他的双手，宣布在这种情况下是不可能谈论意义的，并因此怎样都行，而是坚持认为，我们应该对这种未知的未来承担起责任。在伦理决策的情况下，这将导致一种难题(133～135 页)：我们必须对我们决策的未来效果承担责任，但是我们却不可能知道那些效果，也不可能等待到能看见它们是什么。我们现在必须作出决策。

我们如何来处理这种难题呢？返回到普遍原理中去等于否认了我们生活于其中的社会系统的复杂性，因此决不是正当的做法。放任自流，就是在推卸我们的责任。对于此系统的前一种趋法是过于僵硬了，而后一种趋法则是过于灵活了。康奈儿的建议(追随德里达，重新以我的术语进行表述)是严肃地采取一种现有的伦理(以及

法律）原理——抵制变化——并且能敏锐地意识到当它们何时不再适用或必须放弃。因此，我们的确在遵循一些原理，就好像它们是普遍规则［康奈尔和德里达运用的术语是“准超验”（quasi-transcendental）］一样，但是当我们每次使用时，就必须重新激发规则的合法性。伦理地行动，意味着不是盲目地追随规则——仅仅进行计算——而是负责任地遵循它们，这也许意味着，规则必须被打破。重要的是要强调，在这些情形下，规则被打破并不意味规则失效。如果规则是某种受到逻辑关系束缚的抽象规则集的一部分，则的确如此。但是，如果是从复杂的关系集中形成了某种准规则，那么这种规则的结构的一部分就有不遵循它的可能。要作出一个负责任的判断——而不论它是法律、科学还是艺术——因此至少便会涉及如下的成分：

- 尊重其他人及其自身价值的差异。
- 尽可能多地收集起该问题的信息，尽管不可能收集起全部的信息。
- 尽可能多地考虑到该判断的后果，尽管不可能考虑到所有的后果。
- 保证一旦其缺陷已经清楚就可能修改判断，无论是在特定的还是一般的情形下。

我必须承认，对于康奈尔的工作的扼要讨论并没有充分地评价她提出的许多问题。伦理、公正和法律的问题——在个人、国家和国际的层面上——仍然是某种需要面对的最具挑战性、最为重要的问题。我希望表明，从后结构视野进行解释的复杂性理论，可能有助于我们以更好的手段趋进这些问题。

注 释：

1．在前一章中，我们碰到“群体思维”的概念。作为一种等价物——在利奥塔尔的倾向于差异敏感性的主张的语境中——我们可以创造一个术语叫“差异思维”，暗指这种方式，即群体的成员只可能是以他们相互的差异而区分。

选用“形似性”这个词也是有显著意义的。它通常被用来设计谬误推理，标志某种不合逻辑的东西。这个词本义是指“除了”的逻辑，而利奥塔尔使用它来表明，当处理偶发事物复杂性的丰富性和矛盾时，逻辑描述是不适用的。许多故事，甚至是充满着矛盾的故事，也可以讲述单个事件或现象。利奥塔尔(1984：60)将形似性与“创新”(innovation)区分开来，他认为，这仍然处于系统的命令之下，因为它被用来改进后者的效率。形似性是“在知识语用论中上演的电影”，其后果不可能预先决定。

2．在此利奥塔尔明确地针对鲍德里亚，鲍德里亚对于后现代社会的分析常常有非常否定的味道。

3．在她关于社会(或伦理)关系的分析中，康奈尔(Cornell 1992)也意识到对称关系的必然性。为了这种不当利用的可能性最小化，她坚持(沿着德里达对于列维纳斯的阅读)“现象学的对称”(85 页)。此论据在于，系统中的任何要素都不是先验地优先的。在它们的存在中，所有的要素都是平等的。对于获得关于复杂系统中要素的理解并且不与这种理解冲突，这是重要的一点。不过，它并不能对于抵制不当利用提供任何保证，也不能在我们遇上了的时候真正帮助我们澄清不当利用。问题在于，在很大的程度上，网络中的节点(以及因此在社会系统中的个体)是由其关系所构成，而不是由节点本身的某种现象学方面所构成。我们必须赋予每一个体的现象学等价性，但是，所有之中最明显的情形是，要确定某人是否已经受到了侵犯，我们必须开始解开社会关系之网。这些关系是偶发的，因此每次都是不同的。

4．必须注意到，利奥塔尔对此的评价中是有着某种雄心的。尽管他表明，“腐蚀的潜势”是内在于传统的科学话语的，但他并未感到在此过程中有某种重要的东西丢失了，某种特定的“思辨性”研究只有在属于传统大学的“系科组织”的支持下才是可能的(Lyotard 1984：39，52)。

5．这点形成了利奥塔尔拒斥解释学的基础。解释学话语形成的前提是“真正的知识……由报告陈述所组成，这样的陈述结合到保证其合法性的主题的元叙事之中”。这种前提“保证了存在着可以知道的意义，从而将赋予历史(特别是学识的历史)以合法性”(Lyotard 1984：35)。

6．不过，Rouse (1991b：609)的确认为自己是被她称作“欢乐后现代”(cheerfully postmordern)群体的一员。

7．丘奇兰德的哲学风格常常被看作“消除的唯物论”(eliminative materialism)(参见 P. M. Churchland 1989：1～122)。

8．在主张简单性的超验价值中，丘奇兰德选择了使用最小网络(181 页)。我在第 5 章中提出了不同于他们的针对性看法。

9．我分析了这篇冠名“种族主义的最后一言”(Racism's Last Word)(Derrida 1985)的论文，文章待发表(Gilliers 1998)。

第 8 章　后记　理解复杂性

本书的目的在于，激起人们对复杂性主题的跨学科讨论。我认为，传统的科学方法和分析哲学，对于复杂系统的动力学都不够敏感。因此，我已经对分析方法[把事物割裂(carving things up)]、演绎逻辑、原子论、基于形式规则的语法、封闭的算法和符号表征进行了批判。这些趋法的缺点，在对于信息理论、形式符号系统、塞尔的“中文房间”和福多尔精神表征主义(mental representationalism)的分析中已加以指出。作为一种可替代的选择，我提议采取一种联结论趋法，主张这种趋法对于复杂性是内在地更为敏感。关注的焦点是许多相互关联的、类似的要素的集体行为，这些要素并不因其本身而具有(原子论的)意义，而是通过网络中的非线性、不对称关系的复杂集合而获得了意义。这些网络的重要特点，包括分布性、自组织和无中心控制的局域信息运行。我还表明了，这些模型已经得到了实际应用，例如应用于模式识别问题，而且，将它们理解为复杂系统有助于我们去改进它们的实际性能。在整本书中，我力图将哲学话语与科学

话语交织起来。已经表明的思想是，不仅仅是哲学的考虑将如何有益于科学实践，而且反之亦然。特别在第 7 章中，我力图表明对于复杂性的实践理解，希望能对后现代哲学的某些关键领域有所贡献。

就这些复杂性模型的应用而言，这个结论不过是一个开端。必须清楚地意识到，这里发展的对于复杂性的一般理解，并不对任何特定的复杂系统提供某种完整的描述。如果我们希望(例如)将活细胞作为复杂系统进行分析，这里提出的思想就只能提供一个框架。此框架中必须填充有关此特定情形的(生物化学的)细节。类似地，我们已经从复杂系统刻画了语言的特征，但是这必须被看作是非常初步的尝试。要建立起有用的、大范围的联结论语言模型，以及它们在机器中的执行——这应该被证明是可能的——还有大量的艰辛工作要做。我的希望是，这种研究对于解决这些困难的复杂工作，可能提供一定的理论方向。

最后，我非常清楚地意识到，我可能犯下了述行谬误(performative fallacy)，即试图去落实我那说得到而做不到的主张。在这种情况下，问题便将是致力于建立一种理论，坚持激进的条件论而又声称其具有普遍有效性。这与德里达(1981：26)主张“无论哪里都只有差异和痕迹的痕迹”时所面对的问题是完全一样的。要坚持不应该对这种述行谬误负责则使得保持进行任何根本性批判变得极为困难。这里要求在分析的见解与后现代的见解之间作出截然的划分。但正如我们所看见的，在一方面作为决定性论据的东西，却在另一方面成了对于真正问题的逃避。

也许，作为对这种要求的一种真正的后现代反响，便是干脆放弃它。然而，当哈贝马斯对他作出了这种指责时，德里达的反应是相当强烈的[1]，也许，我也应该冒着风险来作出反应。我不可能从论据本身的

基础来提出反驳——在那种情况下它的复杂程度很低(show stopper)。然而，我的确认为，人们应该考虑到被指责为述行谬误的理论之内容。正如德里达的痕迹理论，这里提呈的复杂性的特点是非常稀少的；它主张的东西很少。它只是在非常低的层次上对复杂系统的结构进行了一般性描述。这意味着，较高层次或涌现性质在这样的理论本身中没有任何作用；它们没有“较高”的重要性。正如德里达的语言模型仅仅是由痕迹结合在一起一样，复杂系统也仅是由局域的相互作用结合在一起的。模型并不试图去说明这些相互作用的效果。

这种论点可以在这里最后一次参照有关神经网络训练的类似问题加以清晰说明。神经网络，在其解决问题的方式上被认为是非算法的，但是网络是以学习“算法”而进行训练的。不过，这种学习算法对于所论特定问题的结构却是什么也不知道。同样的“算法”可被用来改变应用于许多问题的网络中的权重。但是，它在如此低的层次上运行，以致其本身是没有内容的，而只有在特定的、条件性的应用中才被赋予了内容。类似地，这里提供的对复杂性的描述，并不假装是一般的，而是低层次的。它并不假装提供了对于任何特定复杂系统的某种精确的、详细的描述。事实上，正因为这种动力学出现在低层次上，这样的精确描述仍然是可疑的。

这一点将肯定不会消除对此问题承担实施了述行谬误的责任或其他的批评，但愿它传达了本书中提呈的精神中的某些东西：一种开放性、暂时性和风险性。

注　释：

1. 关于这种反响可参见关于有限公司(Limited Inc.)的“后记”(Derrida 1988)。

参 考 文 献

Aleksander, I. and Taylor, J. (eds) (1992) *Artificial Neural Networks, 2: Proceedings of the 1992 International Conference on Artificial Neural Networks (ICANN-92)*. Amsterdam: North-Holland.

Argyros, A. J. (1991) *A Blessed Rage for Order: Deconstruction, Evolution, and Chaos*. Ann Arbor: University of Michigan Press.

Austin, J. L. (1980) *How to Do Things with Words*. 2nd edn. Oxford: Oxford University Press.

Axelrod, R. (1984) *The Evolution of Cooperation*. New York: Basic Books.

Bak, P. and Chen, K. (1991) Self-organized criticality. *Scientific American*, January, pp. 26~33.

Bass, A. (1984) The double game: An introduction. In J.H. Smith

and W. Kerrigan (eds), *Taking Chances: Derrida, Psychoanalysis and Literature*. Baltimore: Johns Hopkins University Press, pp. 66~85.

Baudrillard, J. (1984) The precession of simulacra. In B. Wallis (ed.), *Art after Modernism*. New York: New Museum of Contemporary Art, pp. 253~281.

——(1988) *The Ecstasy of Communication*. New York: Semiotext (e).

Bauman, Z. (1992) *Intimations of Postmodernity*. London: Routledge.

——(1993) *Postmodern Ethics*. Oxford: Blackwell.

Bechtel, W. (1987) Connectionism and philosophy of mind: An overview. *Southern Journal of Philosophy*, Vol. XXVI (Supplement), pp. 17~41.

Bierwisch, M. (1990) Perspectives on mind, brain and language: Linguistics as a cognitive science or touring the Chinese Room again. In A. Burkhardt (ed.), *Speech Acts, Meaning and Intentions*. Berlin: Walter de Gruyter, pp. 291~428.

Blackwell, R. J. (1976) A structuralist account of scientific theories. *International Philosophical Quarterly*, Vol. 16, pp. 263~274.

Brooks, R. A. (1991) Intelligence without representation. *Artificial Intelligence*, Vol. 47, pp. 139~159.

Burkhardt, A. (ed.) (1990) *Speech Acts, Meaning and Intention: Critical Approaches to the Philosophy of John R. Searle*.

Berlin: Walter de Gruyter.

Chaitin, G. L. J. (1975) Randomness and mathematical proof. *Scientific American*, May, pp. 47～52.

——(1987) *Algorithmic Information Theory*. Cambridge: Cambridge University Press.

Chandrasekaran, B., Goel, A. K. and Allemang, D. (1988) Connectionism and information processing abstractions. *AI Magazine*, Winter, pp. 24～34.

Changeaux, J. -P., Heidemann, T. and Patte, P. (1984) Learning by selection. In P. Marler and H. S. Terrace (eds), *The Biology of Learning*. Berlin: Springer-Verlag, pp. 115～133.

Chomsky, N. (1957) *Syntactic Structures*. The Hague: Mouton.

——(1972) *Language and Mind*. Enlarged edn. New York: Harcourt Brace Jovanovich, Inc.

——(1980) *Rules and Representations*. New York: Columbia University Press.

Churchland, P. M. (1984) *Matter and Consciousness*. Cambridge, Mass.: MIT Press.

——(1989) *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, Mass.: MIT Press.

Churchland, P. M. and Churchland, P. S. (1990) Could a machine think? *Scientific American*, January, pp. 26～31.

Churchland, P. S. (1986) *Neurophilosophy. Towards a Unified Science of the Mind/Brain*. Cambridge, Mass.: MIT Press.

Churchland, P. S. and Sejnowski, T. J. (1992) *The Computational*

Brain. Cambridge, Mass.: MIT Press.

Cilliers, F. P. (1990) The brain, the mental apparatus and the text: A post-structural neuropsychology. *South African Journal of Philosophy*, Vol. 9, No. 1 (February), pp. 1~8.

——(1998) On Derrida and apartheid. *South African Journal of Philosophy*, Vol. 17, No. 1. Forthcoming.

Clapin, H. (1991) Connectionism isn't magic. *Minds and Machines*, Vol. 1, pp. 167~184.

Clark, A. and Lutz, R. (1992) *Connectionism in Context*. London: Springer-Verlag.

Clark, T. (1985) Computers as universal mimics: Derrida's question of mimesis and the status of 'artificial intelligence'. *Philosophy Today*, Winter, pp. 303~318.

Cornell, D. (1992) *The Philosophy of the Limit*. London: Routledge.

Culler, J. (1976) *Saussure*. London: Fontana Press.

——(1983) *On Deconstruction*. London: Routledge and Kegan Paul.

Cummins, R. (1991) Form interpretation, and the uniqueness of content: Response to Morris (1991). *Minds and Machines*, Vol. 1, pp. 31~42.

Denning, P. J. (1990) Is thinking computable? *American Scientist*, Vol. 78 (March-April), pp. 100~102.

Derrida, J. (1973) *Speech and Phenomena, and Other Essays on Husserl's Theory of Signs*. Evanston, Ill.: Northwestern University Press.

——(1976) *Of Grammatology*. Baltimore: Johns Hopkins University Press.

——(1978) *Writing and Difference*. Chicago: University of Chicago Press.

——(1981) *Positions*. Chicago: University of Chicago Press.

——(1982) *Margins of Philosophy*. Brighton: Harvester Press.

——(1983) The principle of reason: The university in the eyes of its pupils. *Diacritics*, Fall, pp. 3~20.

——(1985) Racism's last word. *Critical Inquiry*, Autumn, pp. 290~292.

——(1988) *Limited Inc*. Evanston, Ill.: Northwestern University Press.

Derthick, M. (1990) Review of S. Pinker and J. Mehler, 1988: *Connections and Symbols* (MIT Press, Cambridge, MA). *Artificial Intelligence*, 43 (1990), pp. 251~265.

Dreyfus, H. L. and Dreyfus, S. E. (1986) *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. New York: Free Press.

Eco, U. (1987) *Travels in Hyperreality*. London: Picador.

Edelman, G. M. (1987) *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books.

Feyerabend, P. (1975) *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: Verso.

Flanagan, S. (1990) *Hildegard of Bingen. A Visionary Life*. London: Routledge.

Fodor, J. A. (1975) *The Language of Thought*. New York: Thomas Crowell.

——(1981) *Representations*. Cambridge, Mass.: MIT Press.

Fodor, J. A. and Pylyshyn, Z. W. (1988) Connectionism and cognitive architecture: A critical analysis. In S. Pinker and J. Mehler (eds), *Connections and Symbols*. Cambridge, Mass.: MIT Press, pp. 3～71.

Freud, S. (1925) A note on the 'Mystic Writing-Pad'. In *Standard Edition*, Vol. 19. London: The Hogarth Press, pp. 225～232.

——(1950) *Project for a Scientific Psychology*. In *Standard Edition*. Vol. 1. London: The Hogarth Press, pp. 281～397.

Garfinkel, A. (1987) The slime mold *dyctyostelium* as a model of self-organization in social systems. In F. E. Yates (ed.), *Self-Organizing Systems: The Emergence of Order*. New York: Plenum Press, pp. 181～212.

Gibson, J. J. (1979) *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Globus, G. G. (1991) Deconstructing the Chinese Room. *Journal of Mind and Behaviour*, Vol. 12, No. 3, pp. 377～391.

Gödel, K. (1962) *On Formally Undecidable Problems*. New York: Basic Books.

Goodman, N. D. (1991) Review of Putnam (1988). *Minds and Machines*, Vol. 1, pp. 117～119.

Grossberg, S. (1982) *Studies of Mind and Brain*. Dordrecht: D. Reidel.

——(ed.) (1987) *The Adaptive Brain*. 2 vols. Amsterdam: North-Holland.

—— (ed.) (1988) *Neural Networks and Natural Intelligence*. Cambridge, Mass.: MIT Press.

Guenthner, F., Lehman, H. and Schönfeld, W. (1986) A theory for the representation of knowledge. *IBM Journal of Research and Development*, Vol. 30, No. 1 (January), pp. 39~56.

Harland, R. (1987) *Superstructuralism: The Philosophy of Structuralism and Post-Structuralism*. London: Methuen.

Harnad, S. (1989) Minds, machines and Searle. *Journal for Experimental and Theoretical Artificial Intelligence*, Vol. 1, pp. 5~25.

Haugeland, J. (1981) *Mind Design: Philosophy, Psychology, Artificial Intelligence*. Montgomery, Ala.: Bradford Books.

—— (1985) *Artificial Intelligence: The Very Idea*. Cambridge, Mass.: MIT Press.

Hebb, D. O. (1949) *The Organization of Behaviour*. New York: Wiley.

Hecht-Nielsen, R. (1992) The munificence of high dimensionality. In I. Aleksander and J. Taylor (eds), *Artificial Neural Networks, 2: Proceedings of the 1992 International Conference on Artificial Neural Networks (ICANN-92)*. Amsterdam: North-Holland, pp. 1017~1030.

Hesse, M. (1988) Theories, family resemblances and analogy. In D. H. Helman (ed.), *Analogical Reasoning*. Dordrecht: Kluwer

Academic Press, pp. 317～340.

——(1992) Models, metaphors and truth. In F. R. Akkersmit and J. J. A. Mooij (eds), *Knowledge and Language III*. Dordrecht: Kluwer Academic Press, pp. 49～66.

Hirose, Y., Koichi, Y. and Hijiya, S. (1991) Back-propagation algorithm which varies the number of hidden units. *Neural Networks*, Vol. 4, No. 1, pp. 61～66.

Hofstadter, D. R. (1980) *Gödel, Escher, Bach: An Eternal Golden Braid*. Harmondsworth: Penguin Books.

Hofstadter, D. R. and Dennett, D. C. (eds) (1982) *The Mind's I: Fantasies and Reflections on Self and Soul*. Harmondsworth: Penguin Books.

Horgan, T. and Tienson, J. (1987) Settling into a new paradigm. *Southern Journal of Philosophy*, Vol. XXVI (Supplement), pp. 97～113.

Hornick, K., Stinchcombe, M. and White, H. (1989) Multilayer feedforward networks are universal approximators. *Neural Networks*, Vol. 2, No. 5, pp. 359～366.

Jen, E. (ed.) (1990) *1989 Lectures in Complex Systems*. Redwood City, Calif.: Addison-Wesley.

Judd, S. (1992) Why are neural networks so wide? In I. Aleksander and J. Taylor (eds), *Artificial Neural Networks, 2: Proceedings of the 1992 International Conference on Artificial Neural Networks (ICANN -92)*. Amsterdam: North-Holland, pp. 45～52.

Katz, B. F. and Dorfman, M. H. (1992) The neural dynamics of conversational coherence. In A. Clark and R. Lutz (eds), *Connectionism in Context*. London: Springer-Verlag, pp. 167~181.

Kauffman, S. A. (1991) Antichaos and adaptation. *Scientific American*, August, pp. 64~70.

——(1993) *The Origins of Order: Self-Organisation and Selection in Evolution*. New York: Oxford University Press.

——(1995) *At Home in the Universe: The Search for Laws of Complexity*. London: Viking Press.

Kohonen, T. (1988) *Self-Organization and Associative Memory*. 2nd edn. Berlin: Springer-Verlag.

Kolen. J. F. and Goel, A. K. (1991) Learning in parallel distributed processing networks: Computational complexity and information content. *IEEE Transactions on Systems Man and Cybernetics*, Vol. 21, No. 2 (March-April), pp. 359~367.

Krohn, W. and Küppers, G. (1989) Self-organization: A new approach to evolutionary epistemology. In K. Hahlweg and C. A. Hooker (eds), *Issues in Evolutionary Epistemology*. Albany, NY: State University of New York Press. pp. 151~170.

Lawson, H. (1985) *Reflexivity. The Post-Modern Predicament*. London: Hutchinson.

Lewin, R. (1993) *Complexity: Life on the Edge of Chaos*. London: Phoenix.

Lloyd, D. E. (1989) *Simple Minds*. Cambridge, Mass.: MIT Press.

Luhmann, N. (1985) *A Sociological Theory of Law*. London: Routledge and Kegan Paul.

Lyotard, J. -F. (1984) *The Postmodern Condition: A Report on Knowledge*. Manchester: Manchester University Press.

McCulloch, W. S. and Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115~133.

Marletti, D. (1990) Congruence and modulation in propositional attitudes. In A. Burkhardt (ed.), *Speech Acts, Meaning and Intentions*. Berlin: Walter de Gruyter, pp. 279~299.

Minsky, M. and Papert, S. (1969) *Perceptrons*. Cambridge, Mass.: MIT Press.

Morris, M. (1991) Why there are no mental representations. *Minds and Machines*, Vol. 1, pp. 1~30.

——(1992) Beyond interpretation. Reply to Cummins (1991). *Minds and Machines*, Vol. 2, pp. 85~95.

Mountcastle, V. B. (1978) An organizing principle for cerebral function: The unit module and the distributed system. In G. M. Edelman and V. B. Mountcastle, *The Mindful Brain*. Cambridge, Mass.: MIT Press, pp. 7~50.

Münch, D. (1990) Minds, brains and cognitive science. In A. Burkhardt (ed.), *Speech Acts, Meaning and Intention*. Berlin: Walter de Gruyter, pp. 367~390.

Nicolis, G. and Prigogine, I. (1989) *Exploring Complexity*. New York: Freeman and Co.

Norris, C. (1992) *Deconstruction and the Interest of Theory*. London: Leicester University Press.

Parushnikova, A. (1992) Is a postmodern philosophy of science possible? *Studies in the History and Philosophy of Science*, Vol. 23, No. 1, pp. 21~37.

Pattee, H. H. (1987) Instabilities and information in biological self-organization. In F. E. Yates (ed.), *Self-Organization: The Emergence of Order*. New York: Plenum Press, pp. 325~338.

Penrose, R. (1989) *The Emperor's New Mind*. Oxford: Oxford University Press.

Pribram, K. H. and Gill, M. M. (1976) *Freud's 'Project' Reassessed*. London: Hutchinson.

Prigogine, I. and Stengers, I. (1984) *Order out of Chaos: Man's New Dialogue with Nature*. London: Heinemann.

Putnam, H. (1988) *Representation and Reality*. Cambridge, Mass: MIT Press.

Rapaport, W. J. (1986) Philosophy, artificial intelligence, and the Chinese Room argument. *Abacus*, Vol. 3, No. 4, pp. 7~17.

Rorty, R. (1980) *Philosophy and the Mirror of Nature*. Oxford: Blackwell.

Rosenblatt, F. (1958) The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, pp. 386~408.

Rosenfeld, I. (1988) *The Invention of Memory: A New View of the*

Brain. New York: Basic Books.

Rouse. J. (1990) The narrative reconstruction of science. *Inquiry*, Vol. 33, No. 1, pp. 179~196.

——(1991a) Philosophy of science and the persistent narratives of modernity. *Studies in History and Philosophy of Science*, Vol. 22, No. 1, pp. 141~162.

——(1991b) The politics of postmodern philosophy of science. *Philosophy of Science*, Vol. 58, pp. 607~627.

——(1996) *Engaging Science: How to Understand Its Practices Philosophically*. Ithaca, NY: Cornell University Press.

Rumelhart, D. E. and McClelland, J. L. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, 2 vols. Cambridge, Mass.: MIT Press.

Saussure, F. de (1974) *Course in General Linguistics*. London: Fontana.

Schalkwyk, D. (1991) The meaning of wor(l)d: Value, sense, and reference in *Not Saussure* by Raymond Tallis. *Journal of Literary Studies*, Vol. 7, No. 3/4 (December), pp. 193~216.

Schrag, C. O. (1992) *The Resources of Rationality: A Response to the Postmodern Challenge*. Bloomington: Indiana University Press.

Searle, J. R. (1969) *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

——(1977) Reiterating the differences: A reply to Derrida. *Glyph*, Vol. 2, pp. 198~208.

——(1979) Intentionality and the use of language. In A. Margalit (ed.), *Meaning and Use*. Dordrecht: Kluwer Academic Press, pp. 181~197.

——(1980) Minds, brains, and programs. *Behavioral and Brain Sciences*, Vol. 3, pp. 417 ~ 457; Continuing Commentary in Vol. 5 (1982), pp. 338~348; Vol. 8 (1985), pp. 741~743.

——(1982) The Chinese Room revisited. *Behavioral and Brain Sciences*, Vol. 5, pp. 345~348.

——(1983) *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

——(1984) *Minds, Brains and Science*. Harmondsworth: Penguin Books.

——(1985) Patterns, symbols, and understanding. *Behavioral and Brain Sciences*, Vol. 8, pp. 742~743.

——(1990) Is the brain's mind a computer program? *Scientific American*, January, pp. 20~25.

Serra, R. and Zanarini, G. (1990) *Complex Systems and Cognitive Processes*. Berlin: Springer-Verlag.

Shannon, C. E. (1948) A mathematical theory of communication. *Bell System Technical Journal*, Vol. 27, pp. 379~423.

——(1949) Communication in the presence of noise. *Proceedings of the Institution of Radio Engineers*, Vol. 37 (Jan.), pp. 10~21.

Sietsma, J. and Dow, R. J. F. (1991) Creating artificial neural networks that generalize. *Neural Networks*, Vol. 4, No. 1, pp. 67~79.

Smolensky, P. (1987) The constituent structure of connectionist mental states: A reply to Fodor and Pylyshyn. *Southern Journal of Philosophy*, Vol. XXVI (Supplement), pp. 137～161.

Staufer, D. (1987) Random Boolean networks: Analogy with percolation. *Philosophical Magazine B*, Vol. 56, No. 6, pp. 901～916.

Sterelny, K. (1990) *The Representational Theory of Mind*. Oxford: Blackwell.

Stofberg, J. A. (1988) Objectivity and the sociology of science. *South African Journal of Philosophy*, Vol. 7, No. 4, pp. 213～225.

Toffoli, T. and Margolus, N. (1987) *Cellular Automata Machines: A New Environment for Modelling*. Cambridge, Mass: MIT Press.

Touretzky, D. S. and Hinton, G. E. (1985) Symbols among the neurons: Details of a connectionist inference architecture. *Proceedings of the 9th International Joint Conference on AI*, Los Angeles, Calif., 18～23 August, pp. 238～243.

Turing, A. (1936) On computable numbers, with an application to the 'Entscheidungsproblem'. *Proceedings of the London Mathematic Society*, Series 2, 42, pp. 230～265.

——(1950) Computing machinery and intelligence. *Mind*, Vol. LIX, No. 236. Reprinted in D. R. Hofstadter and D. C. Dennett (eds), *The Mind's I: Fantasies and Reflections on Self and Soul*. Harmondsworth: Penguin Books, 1982, pp. 53～67.

Von der Malzburg, C. (1987) Ordered retinotectal projections and brain organization. In F. E. Yates (ed.), *Self-Organizing Systems: The Emergence of Order*. New York: Plenum Press, pp. 265～277.

Von Foerster, H. (1985) Entdecken oder Erfinden—Wie lasst sich verstehen verstehen? In H. Gumin and A. Mohler (eds), *Einführung in den Konstruktivissmus*. Munich: Schriften der Carl-Friedrick-von-Siemens-Stiftung, pp. 27～68.

Wilden, A. (1984) *System and Structure: Essays in Communication and Exchange*, 2nd edn. London: Tavistock Publications Ltd.

——(1987) *The Rules are No Game: The Strategy of Communication*. London: Routledge and Kegan Paul.

译　后　记

我们刚刚成功地申办了第 13 届国际逻辑学、方法论和科学哲学大会(The 13^{th} International Congress of Logic, Methodology and Philosophy of Science)，大会将于 2007 年在清华大学召开，这是大会首次在我国举行，也是首次在亚太国家举行。为了进一步拓展科技哲学研究的国际视野，促进对话，我们与上海科技教育出版社的合作也就加快了步伐。利用今年春节在内的整个寒假及其前后的一段时间，我几乎把自己关闭起来，终于译出了初稿并校对了一遍。

本书讨论的是“复杂性与后现代主义”，而后现代主义于我来说是一个非常陌生的领域，在翻译中也就遇到了不少的困难。幸好，在译完初稿后，见到了文兵、宋朝龙、杨生平于 2003 年下半年译成的交流稿(北京大学现代科学与哲学研究中心组织编印)，从而得以参照该译本对一些翻译包括名词的确定进行了推敲。特别是，清华大学科学技术与社会研究中心(STS 中心)的张成岗博士已经翻译出版了《后现代伦理学》一书，蒋劲松博士在后现代哲学上素有研究，吴彤

教授在复杂性的哲学问题研究上颇有造诣，他们分别阅读了基本成形的全部或部分译稿，然后一起就其中的若干翻译问题进行了讨论。

随后，我利用“五一”长假及其前后的一段时间进行了又一次校订。然而种种事务缠身，最后一直拖到8月底才得以定稿，并交付出版社。即使如此，不妥乃至误译仍然难以避免，有待改进之处更不在少数，敬请各位同仁给予指正。

曾国屏

于清华大学STS中心

sts001@tsinghua.edu.cn

2004年9月